



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 106297

TO: Phuong Bui
Location: 9a09 / 9e12
Tu sday, October 21, 2003
Art Unit: 1638
Phone: 305-1996
Serial Number: 09 / 868546

From: Jan Delaval
Location: Biotech-Chem Library
CM1-1E07
Phone: 308-4498

jan.delaval@uspto.gov

Search Notes

Jan Delaval
Reference Librarian
Biotechnology & Chemical Library
CM1 1E07 - 703-308-4498
jan.delaval@uspto.gov

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2003, 01:49:44 ; Search time 6721 Seconds
(without alignments)
11315.429 Million cell updates/sec

Title: US-09-868-546A-1
Perfect score: 1859
Sequence: 1 Gaaacactgacagacgca.....aataaatttcttcttc 1859

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_man.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_higo_hum.*
- 40: em_higo_mus.*
- 41: em_higo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	777.4	41.8	1820	8	AB022732	AB022732 Glycyrhri
2	769.4	41.4	1806	8	AB001379	AB001379 Glycyrhri
3	735.4	39.6	1668	8	CAR012581	AJ012581 Cicer ari
4	720.2	38.7	1500	8	AB025016	AB025016 Lotus jap
5	716.6	38.5	1781	8	CAR238439	AJ238439 Cicer ari
6	499.8	26.9	1711	8	CAR239051	AJ239051 Cicer ari
7	415.8	22.4	104321	8	AP006142	AP006142 Lotus jap
8	394.8	21.2	1693	8	BT002874	BT002874 Arabidops
9	394.4	21.2	1519	8	BT004449	BT004449 Arabidops
10	393.8	21.2	1488	6	AX506493	AX506493 Sequence
11	393.6	21.2	1561	8	AY113869	AY113869 Arabidops
12	393.6	21.2	1612	8	AY050849	AY050849 Arabidops
13	386.6	20.8	1494	6	AX507583	AX507583 Sequence
14	386.6	20.8	1494	6	AX590025	AX590025 Sequence
15	386.6	20.8	1494	6	AX651453	AX651453 Sequence
16	386.6	20.8	1525	8	BT000090	BT000090 Arabidops
17	386.6	20.8	1620	8	AY065192	AY065192 Arabidops
18	384.6	20.7	1656	8	AY039844	AY039844 Arabidops
19	384.6	20.7	1656	8	AY087256	AY087256 Arabidops
20	384	20.7	1652	8	D78607	D78607 Arabidops
21	383.8	20.6	1479	6	AX507477	AX507477 Sequence
22	383.2	20.6	1503	6	AX506778	AX506778 Sequence
23	383.2	20.6	1503	6	AX651658	AX651658 Sequence
24	381	20.5	1653	8	AY086486	AY086486 Arabidops
25	378.4	20.4	1503	6	AY074823	AY074823 Arabidops
26	377	20.3	1503	6	AX505366	AX505366 Sequence
27	377	20.3	1664	8	AY087977	AY087977 Arabidops
28	366.4	19.7	1508	6	BT000794	BT000794 Arabidops
29	361.8	19.5	1509	6	AX412388	AX412388 Sequence
30	361.8	19.5	1509	6	AX506140	AX506140 Sequence
31	361.8	19.5	1509	6	AX651784	AX651784 Sequence
32	361.8	19.5	1509	8	BT001045	BT001045 Arabidops
33	361.8	19.5	1587	8	AY093766	AY093766 Arabidops
34	360.2	19.4	1580	8	D78606	D78606 Arabidops
35	357	19.2	1507	8	AY096511	AY096511 Arabidops
36	357	19.2	1740	8	AY065209	AY065209 Arabidops
37	354	19.0	1578	8	AY084273	AY084273 Arabidops
38	351.4	18.9	1674	6	BD081607	BD081607 Plant fat
39	351.4	18.9	1674	8	HTCYP81C	AJ000477 Helianthu
40	350.6	18.9	1557	6	AX507264	AX507264 Sequence
41	350.6	18.9	1557	6	AX652005	AX652005 Sequence
42	349.2	18.8	1596	8	AY054214	AY054214 Arabidops
43	345	18.6	976	8	CAR249800	AJ249800 Cicer ari
44	344.8	18.5	1719	8	HTCYP81L	AJ000478 Helianthu
45	337.4	18.1	1632	6	AX412401	AX412401 Sequence

ALIGNMENTS

RESULT 1
AB022732

LOCUS AB022732
DEFINITION Glycyrhiza echinata CYP Ge-31 mRNA for cytochrome P450, complete cds, linear PLN 21-APR-2000

ACCESSION AB022732
VERSION AB022732.1 GI:4200043
KEYWORDS cytochrome P450.
SOURCE Glycyrhiza echinata
ORGANISM Glycyrhiza echinata

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Galegeae;
Glycyrhiza
REFERENCE 1 (sites)

AUTHORS Nakamura, K., Akashi, T., Aoki, T., Kawaguchi, K. and Ayabe, S.
TITLE Induction of isoflavonoid and retrochalcone branches of the flavonoid pathway in cultured Glycyrrhiza echinata cells treated with yeast extract
JOURNAL Biotechnol. Biochem. 63 (9), 1618-1620 (1999)
MEDLINE 20008263
PUBMED 10540749
REFERENCE 2 (bases 1 to 1820)
AUTHORS Akashi, T., Aoki, T. and Ayabe, S.
TITLE Direct Submission
JOURNAL Submitted (23-JAN-1999) Shin-ichi Ayabe, Nihon University, Department of Applied Biological Science, Kameino 1866, Fujisawa, Kanagawa 252-8510, Japan [E-mail: ayabe@brs.nihon-u.ac.jp, Tel: 81-466-84-3703, Fax: 81-466-80-1141]
FEATURES Location/Qualifiers
 1..1820
 /organism="Glycyrrhiza echinata"
 /mol_type="mRNA"
 /db_xref="taxon:46348"
 /cell_type="cultured cells"
 /clone_lib="lambda Zap II"
 /note="Cultivar maintained at Medicinal Plant Garden, School of Pharmaceutical Sciences, Kitasato University"
 1..1820
 /gene="CYP Ge-31"
 91..1590
 /gene="CYP Ge-31"
 /codon_start=1
 /product="cytochrome P450"
 /protein_id="BAA74465.1"
 /db_xref="GI:4200044"
 /translation="MDILSLSYSVFLAFIFNIVIRARKFNKLPGPSLPFIIGN LHLKRLPHRTFKGLSEKYGHVILMFGSLVVVSSASFPQCFTKNDVILANRPRF LSGKYIFVNYTLGASTSGEHWNLRIITDALVSLNHRINSFSGRDETORLITLADSSTNFAEILSVRFYDMFTNIMRISGKRYGDCDMSDLOEASQFRDMVSELQ LSGANNKDFMPLRDLFENLEKRLKDISGTFDAFLRGLIQEHRAKKERANTMIDHL INLDQSPDYTDQIIKGLAMLAGDSSAVTLEWSMLNHLNHPVLRKVKDELDT HVGDRLVDESDPKLSYLNKVINETRLYTPAPLPHSTSDCNETGYKVPQDITV LINAWIHRDPELMEATYTFKPERFEKGELEKLIAPFGMRACPGGLAIRAISMTL ALLIQCPDWKLTNGDKIDMAERDGFITKLVLPMCKSRPVIKVPKQ"

source

gene
 CDS
 519 a 433 c 398 g 470 t
 Query Match 41.8%; Score 777.4; DB 8; Length 1820;
 Best Local Similarity 70.6%; Pred. No. 3.9e-165;
 Matches 1055; Conservative 0; Mismatches 431; Indels 9; Gaps 1;
 75 TCTCATAGCTGCTCTTTCTCGTGTCTATTCCTCGGCGTGAAGTTTGTTCCTCAAGCA 134
 110 TGTCTCTCTCGTCTTCTATCTTCTGCTCTATTTTTCATATTCACATTTGTGATCAGACAA 169
 135 GAAATTCAGAAACATACACAGGTCTCTCTCTCTTCCCATATAGGAAACCTTACC 194
 170 GAAATTCAGAAACCTCCACCGGTCCACCTTCGCTTCCCATATTCGAAACCTCCACC 229
 195 TCCTCGAAGCAGCAATCCACCGTTTCTTCCACGATGTGAAACAGTACGCGCAACGTGG 254
 230 ACCTCAACGCGCCCTCCACGCGACCTTCAGGGGTGTGCGGAAAGTACGCGCATGTGA 289
 255 TTTCCCTCTGTTTCGGTTCAGCTCTGGCGGTGTGTCATCTCTCTCCACAGCATACCAAG 314
 290 TTTCCCTATGTTTGGATCAGCCCTCGTGGTGTGATTTTCTCTCGCATCTGAATTCACAC 349
 315 AATGCTTCACCAACACAGCGTGTGCTTGGCCCAACCGGCTACTTCTCTCTCGGAAAT 374
 350 AGTGTTTCAACCAACAGCGTGTGCTTGGCCCAACAGCGCCGCTTCTCTCGGAAAT 409
 375 ACATCTTCAACCAACACCGTAGGCTCTGCTCCACGCGGAGACACTGGCGCAACC 434
 410 ACATCTTTTACAACTACACCACTTGGGGTCCACATCTCTATGGCGAGCACTGGGCAACC 469
 435 TCCGCGGATCACCGCCCTGGACGTCCTCTCCACGCGGCGGTTCACCTCTCTCTCCGAA 494

BASE COUNT

519 a

433 c

398 g

470 t

ORIGIN

Db 470 TCCTGCTATCATCCGCCCTCGAGCTCTCTTCAAACACCGCATCAACAGCTTCTCCGGAA 529
 Qy 495 TCCGGAGCAGCAGAGCAGAGCTCTGATCGAGAGGTGTGCTGGCCCAAGAACTCCGAACG 554
 Db 530 TCCGAAGGACGAGACACAGAGGCTCATACAGGGCTTCCGATGATCTATCCACCAAC- 588
 Qy 555 AGGAAGAGTTTTCGCGAGTGGAGATTAGTTCGATGTTTCAACGACTTAACTTACCAACA 614
 Db 589 -----TTTGGCGAAATAGAACTCAGTTACAGGTTCTTGACATGACATTCACCAACA 640
 Qy 615 TAATGAGATGATATCGGGAGAGAGGTTTACGGAGAGAGAGTGTAGATGAAGACGTTG 674
 Db 641 TCATGAGATGATCTCCGGAGAGAGGTACTATGCGGAGGACTGTGACATGTTCAGATCTTC 700
 Qy 675 AGGAAGCGAGGAGTTTCAGAGAGACTGTGACAGAAATGTTGGAACCTCATGGGCTTGCGTA 734
 Db 701 AAGAAGCAAGCCAGTTTCAGGACATGCTGTCTGAGTTGTTGCAATGTTCTGGGGCCCAACA 760
 Qy 735 ACAAGGAGATCACTTGCCTTTCCTCAGGTGGTTCGATTTTTCAGATGTTGGAGAGCGCT 794
 Db 761 ACAAGACCGACTTCATGCTCCATTTGCTCAGGTTCTCGACTTTGAAAAACCTGGAGAGAGGC 820
 Qy 795 TAAAGAGTATCATTAAGAGGTACGATTCATCTTGAATGAATCCCTTCATGAGACCGTG 854
 Db 821 TCAAGGATATCAGTGGCAAAACCGATGCGCTTCTTGAGAGGACTCATCCAAAGACGACGCG 880
 Qy 855 CCAGCAATGACCGCCAGAAATTCATGATCATCTCTCTCAAACCTGCAAGAGACCCAGC 914
 Db 881 CCAAGAGGACGTCGAAATACCATGATGATCATCTCTCATCTTCCAAGATCTCACAGC 940
 Qy 915 CTCAGTACTACACTGACCAAAATCATCAAAGGCTTCTCTGGCCATGCTTTTGGTGGAA 974
 Db 941 CTGAGTACTACACCGATCAAATCATCAAAGGCTTCTCTGGCCATGCTCTTGGCGGAA 1000
 Qy 975 CTGACTCATCACTGGACTTTAGAGTGTGATCTATTAATTTTATTTGAATCAACCAGAGG 1034
 Db 1001 CCGACTCGTCTGTCTGTAACCTTTGGAGTGTGCGATGCCAATTTGTTGAACATCCAGAGG 1060
 Qy 1035 TGTGAAGAGGCAAGAGATGAATTCGACACTCAAAGTGGGCAAGACCGCTTGTAAATG 1094
 Db 1061 TATTGAAGAGGTAAAGGATGAATTCGATACACATGTTGGCCAAGATCGCTTGGTAGATG 1120
 Qy 1095 AGTCAGACCTTCCAAAACCTTCCATATCTTAGGAAGATCATCTTGTGAGACACTTAGGTGT 1154
 Db 1121 AGTCAGACCTTCCAAAACCTCTCTTACCTTAAAAACGTTATCAATGAGACGCTTCGGTGT 1180
 Qy 1155 ACCCCCGGCCCCAATCTTAATACCTCATGTGCTTCAAGAGATATTAATTTGAAGAT 1214
 Db 1181 ACACCTCGGCTCGGTGTGTACTACCGACCTCGACTTCAGATGAGTGCAATATTGGAGAT 1240
 Qy 1215 TCAATATCCCAAGACACAAATTTGTGATCATTAATGTTGGGCGCATGCAGAGAGATCCTC 1274
 Db 1241 ACAAGGTTCCAAAGACACGATAGTATTGATTAATGCTGGGCCATTCACAGGAGCCCTG 1300
 Qy 1275 AGTTGTGAATGATGCCATGCTTTAACTCAGAGGTTTGTATGTGGAAGAGAGAGAGA 1334
 Db 1301 AATTGTGAGTGAAGCCACGACTTCAAGCCTCAGAGGTTTGAGAAAAAGGGTGAAGTTG 1360
 Qy 1335 AAAAGTTGTAGCATTTGGCATGGAGAGGCTTGGCCAGGAGACCCATGGCTATGC 1394
 Db 1361 AAAAGTTAAATGCTTTGGATGGAGAGAGGCGATGTCCCGGAGAGGCTTGGCTATTC 1420
 Qy 1395 AAAAGTGTGAGCTTTACTTTGGGATTTGATTCATTAATGTTTGTGATGGAACGAGTAAGT 1454
 Db 1421 GGGCAATTAGCATGACTTTGGCGTATTGATTCATGCTTTTGTGGAAGCTTACAAATG 1480
 Qy 1455 AGGAAAGCTTGTATGACAGAGAACTTTGGATTCACCTTGTCAAGGTTAATTTCCATGG 1514
 Db 1481 GTGATAAGATTGATGCGAGAGAGGAGTGAATTCACCTTGAACAAAATTTGGTTCCCTCAA 1540
 Qy 1515 AGGCCATGTGCAAGGCTCGCCCACTTGCCTCAATTAATTTGAATTTAATTAAT 1569
 Db 1541 AGGCCATGTGTAATCACGTCACAGGTTTTCAGAGGATTAATCAT 1595

RESULT 2
 AB001379
 LOCUS
 DEFINITION
 1806 bp mRNA linear PLN 21-APR-2000
 Glycyrrhiza echinata CYP81E1 mRNA for cytochrome P450, complete
 cds, clone: CYP Ge-3.
 ACCESSION
 AB001379 D89431
 VERSION
 AB001379.1 GI:2443347
 KEYWORDS
 CYP81E1, cytochrome P450.
 SOURCE
 Glycyrrhiza echinata
 ORGANISM
 Glycyrrhiza echinata
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Galegeae;
 Glycyrrhiza.
 REFERENCE
 1 (sites)
 AUTHORS
 Akaishi, T., Aoki, T., Takahashi, T., Kameya, N., Nakamura, I. and
 Ayabe, S.
 TITLE
 Cloning of cytochrome P450 cDNAs from cultured Glycyrrhiza echinata
 L. cells and their transcriptional activation by elicitor-treatment
 Plant Sci. 126, 39-47 (1997)
 REFERENCE
 2 (sites)
 AUTHORS
 Akaishi, T., Aoki, T., Kameya, N., Nakamura, I. and Ayabe, S.
 TITLE
 Two new cytochrome P450 cDNAs (Accession Nos. AB001379 and
 AB001380) from elicitor-induced licorice (Glycyrrhiza echinata L.)
 cells (PGR97-167)
 JOURNAL
 Plant Physiol. 115, 1288 (1997)
 REFERENCE
 3 (sites)
 AUTHORS
 Akaishi, T., Aoki, T. and Ayabe, S.
 TITLE
 CYP81E1, a cytochrome P450 cDNA of licorice (Glycyrrhiza echinata
 L.), encodes isoflavone 2'-hydroxylase
 JOURNAL
 Biochem. Biophys. Res. Commun. 251 (1), 67-70 (1998)
 MEDLINE
 9909223
 PUBMED
 9790908
 REFERENCE
 4 (bases 1 to 1806)
 AUTHORS
 Ayabe, S.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (24-FEB-1997) Shin-ichi Ayabe, Nihon University,
 Department of Applied Biological Science; Kameino 1866, Fujisawa,
 Kanagawa 252-8510, Japan (E-mail: ayabe@bra.nihon-u.ac.jp,
 Tel: 81-466-84-3703, Fax: 81-466-80-1141)
 COMMENT
 On Sep 27, 1997 this sequence version replaced gi:1845558.
 D89431: submitted (20-Nov-1996).
 FEATURES
 source
 1. .1806
 /organism="Glycyrrhiza echinata"
 /mol_type="mRNA"
 /db_xref="taxon:46348"
 /clone="CYP Ge-3"
 /cell_type="suspension-cultured cells"
 /clone_lib="lambda Zap II"
 1. .1806
 /genes="CYP81E1"
 69.1568
 /gene="CYP81E1"
 /codon_start=1
 /product="cytochrome P450"
 /protein_id="BAA22422.1"
 /db_xref="GI:2443348"
 /translation="MEILSLSVFYLALFFINIVIRARFKNLPGPPSLPIGN
 LHLKRLPRTFGUGSEKGVFLMFGRLVVVVSASEFOCFKNDVILANPRF
 LSGKIFYNYITLGSYSYGEHRNRLTALDVLNHRINSFSGIRSDVRLITRLA
 DSSNTNFAEMELSSRLYDMTFNNIMISGKRYGDCDSDLQASOFDMVSLQ
 LSGANNKTFMRLPLRDLFENLEKLDISGKTDAFLRGLIEHRTKKERANTMIDHL
 LNLQDSPEVYDQIIKGLALAMLLAGTDSAVTLWSSNLLNHPVLKKVKRLDT
 HVGQRLVDESILPKLTYLKNVINETRLRYTPAPLLPHSTDECNIGYKVPQITIV
 LINAHWRDPELWTEATFKPERFEKGELEKLFAGMGREACRGEGLAIRAISMTL
 ALLIQCFDWKLLNGKIDLAERDFTLTKLVPLKMKCRSPRVINKFQ"
 515 a 428 c 390 g 473 t

Query Match

41.4%; Score 769.4; DB 8; Length 1806;

Best Local Similarity 70.2%; Pred. No. 2.5e-163;
 Matches 1050; Conservative 0; Mismatches 436; Indels 9; Gaps 1;
 QY 75 TCTCATACGCTGTCCTTTTCTCTGGTTCATTCTCGGCGTGAAGTTTGTTCCTTCCAAAGCA 134
 DB 88 TGTCTACTCTCGTCTTCTATCTTGCTCTATTTTTTCATATTCACATTGTGATCAGAGCA 147
 QY 135 GAAAAATTGAGAAACATACACACAGCTCCTCTCTCTTCCCATAATAGGAAACCTTAAAC 194
 DB 148 GAAAAATTCAAAAACCTCCACCGGCTCCACCTTGCTTCCCATAATCGGAAACCTCCAC 207
 QY 195 TCCTCGAACAGCAATTCACCGTTTCTTCCAGCATGTGAAAACAGTACGGCAACGTGG 254
 DB 208 ACCTCAAAAGCCCTCCACCGCACCTTCAAGGGTGTTCGGAAGTACGGCCATGTGT 267
 QY 255 TTTCCCTCTGGTTGGGTTGCTGCGGTTGCTCTCTCTCTCCACAGCATACCAAG 314
 DB 268 TTTCCCTATGTTGGATCAGCCCTGCTGCTGCTGCTCTCTCTGATCTGATTTCCAAAC 327
 QY 315 AATGCTTCCAAACACAGCTTGCCTTTGGCCAAACCGGTACCTTCTCTCGGAAAT 374
 DB 328 AGTGTTCCTTCAAAAAACGCTGCTGCGCAACAGGCCCGCTTCTCTCGGAAAT 387
 QY 375 ACATCTTCTACAAACAGCAGCTAGCTCTCTCCACGGCAGCACTGGCGCAAC 434
 DB 388 ACATCTTTTCAACTACACCACTTGGGGTCCACATCTCTATGGCGAGCACTGGCGCAAC 447
 QY 435 TCCGCGCATCACCGCCCTGACGCTCTCCACGACGGGTCCACTCTCTCTCGGAA 494
 DB 448 TCCGCTGATCACCGCCCTCAGCTCTTTTCAACCACGGATCAACAGCTTCTCGGAA 507
 QY 495 TCCGAGCAGCAGAGCAAGCGTCTGATGACAGAGTGTGTCGCCCAAGAACTCGCAAG 554
 DB 508 TCCGAAGGACGAGACACAAAGGCTCAACAGAGCTTGGCGATGACTCGTCCACGAAC- 566
 QY 555 AGGAAGAGTTTGGCGAGTGGAGATTAGTTCAGTGTTCACAGCTTAATCTTACAAACA 614
 DB 567 -----TTTGGCGAAATGGAATCTAGCTCTATGACATGACGTTTCAACAACA 618
 QY 615 TAATGAGATGATATCGGGGAAGAGTTTACGGAGAGAGAGTGAAGATGAAGACGTTG 674
 DB 619 TCAATGAGATGATCTCTGGGAAGAGATTATTTATGGGAGAGACTGTGACAGCTCAGATCTCC 678
 QY 675 AGGAAGCAGGAGTTCAGAGAGACTGTGACAGAAATTTGGAACTCATGGCTTGGCTA 734
 DB 679 AAGAAGCAAGCCAGTTTCAAGGACATGGTGTCTGAGTTGCTACAGTTGCTGGGGCAACA 738
 QY 735 ACAAGGAGATCATTTGCTTCTCAGTGGTTCGATTTTTCAGATGTGGAGAGCGCT 794
 DB 739 ACAAGACTGACTTTCATGCTTTCAGGTTTCTCGACTTTTGAACCTTAGAGAGAGGC 798
 QY 795 TAAAGAGTATCAGTAAAGAGTACGATTCATCTTGAATAAGATCTTTCATAGAACCGGTG 854
 DB 799 TCAAGGACATCAGTGGGCAAAACCGATGCTTCTTGAGAGGACTCATTAAGAGACCGCA 858
 QY 855 CCAGCAATGACCGCCAGAAATTCATGATCGATCATCTCTCAAACTCAAGAGACCCAGC 914
 DB 859 CCAAGAGAGCGTGTCAAAATACCATGATAGATCATCTCTCAATCTCCAAGACTCACAGC 918
 QY 915 CTCAGTACTACACTGACCAATCATCAAGGCCCTTGTCTGGCCATGCTTTTGTGGNA 974
 DB 919 CTGAGTACTATACCGATCAAAATCATCAAGGCCCTTGTCTGGCCATGCTTCTTGGGGAA 978
 QY 975 CTGACTCATCAACTGGGACTTTAGAGTGGTCAATATCTAAATTTATTAATCAACCCAGAG 1034
 DB 979 CCGACTCTCTGCTGTTACTTTTGGAGTGTGATGCTCAATTTGTTGAACCATCCAGAG 1038
 QY 1035 TGTGAAAGAGCGCAAGAGATGAATTTGGACACTCAAGTGGGCAAGACCGCTTGTGTAATG 1094
 DB 1039 TATTGAAGAGGTAAAGGATGAATTTGGATACACATGTTGGCCAAGATCGCTTGGTAGATG 1098
 QY 1095 AGTCGACCTTCCAAAACCTTCATATCTTAGAGAGATCATCTTGGAGACACTTAGGTTGT 1154


```

Db      848 AACCTGAGTACTACACTGATCAGATCATCAAGCCCTTGCTTTGGCCATGCTCCTTGGCTG 907
QY      972 GAAGTGAATCACTGAGTCTTGAAGTGGTCAATATCAATATTAATTAATGAATCAACCCAG 1031
Db      908 GAACCGACTCATCTGCTGTAACCTTGAAGTGGTCAATGCTAATTTGTTGAACCATCCAG 967
QY      1032 AGGTGTTGAAGAGGCAAGAGTAATGAGACTCAAGTGGGCAAGACCGCTTGTTAA 1091
Db      968 AGATATTTGAAGAGGTAAGAGGATGAATGAGATTCATGTAGGACCAAGATCGTTGGTAG 1027
QY      1092 ATGAGTCAGACCTTCCAAAACCTTCCATATCTTAGGAGATCATCTTGAACACTTAGCT 1151
Db      1028 ATGATATCAGACCTTCCAAAACCTTACTATCTTAAAATGTCACTATGAGACCTTCGGT 1087
QY      1152 TGTACCCCGCCGCCCCCAATTTCTAATACCTCATGTGCTTCAGAGAATATTACAAATGAAG 1211
Db      1088 TGTGTACTCCCGCTCCTTGTATTATACCACTCAACTTCAGATGATGTCATTTGGAG 1147
QY      1212 GATTCATATCCACGAGACACAATTTGTGATCAATTAATGTTGGGGATGCGAGAGATC 1271
Db      1148 GATATAAAGTACCACGAGACACCATAGTATTGATTAATGATGGGCCATTCACAGATC 1207
QY      1272 CTCAGTTGTGGAATGATGCCACATGCTTTAAACCTGAGAGTGTGATGCGAAGGAGG 1331
Db      1208 CTAAAGTCAGAGTGGAGCCCAAGCTTCAAGCCGAGAGGTTTGACAAAAGGGGAGA 1267
QY      1332 AGAAAAGTTGGTAGCATTTGGCATGGGAAGAGGCTTGGCCAGGAGAACCCATGGCTA 1391
Db      1268 TTGAAAAGGTGATTGCAATTTGGCATGGGAAGAGAGTGTCTCTGGAGAAGCTTTGGCTC 1327
QY      1392 TGCAAGTGTGACCTTACTTTGGGATTTGTTGATTCATGTTTGTGCTGGAACGAGTAA 1451
Db      1328 TTGAAACAAATTAGCATGACTTTGGCAATATTGTTCAATGCTTTGATTTGGAACGATCAA 1387
QY      1452 GTGAGGAAAGCTTTGATATGACAGAGAACAATTTGGATCACTTGTCAAGGTTAATTCAT 1511
Db      1388 GTGATGACAGATTTGATATGGCAGAACGATGGATTGTTTTAACAAATTAATTCAT 1447
QY      1512 TGGAGCCATGTGCAAGGCTGCGCCACTTGCCACTAAAATTTGAATTAATTA 1564
Db      1448 TAAAGGCCATGTGTAAGACTCGTCCGGTGTGTCACCAAGGTTTTCAAGTAATTA 1500

RESULT 4
AB025016 1500 bp mRNA linear PLN 01-FEB-2001
LOCUS    Lotus japonicus mRNA for cytochrome P450, complete cds.
DEFINITION
AB025016 Lotus japonicus
VERSION  AB025016.1 GI:7415995
KEYWORDS cytochrome P450.
SOURCE   Lotus japonicus
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
          Lotus.
REFERENCE
AUTHORS  Shimada,N., Akashi,T., Aoki,T. and Ayabe,S.
TITLE    Induction of isoflavonoid pathway in the model legume Lotus
         japonicus: molecular characterization of enzymes involved in
         phytoalexin biosynthesis
JOURNAL  Plant Sci. 160 (1), 37-47 (2000)
PUBMED   11164575
REFERENCE 2 (bases 1 to 1500)
AUTHORS  Shimada,N., Aoki,T. and Ayabe,S.
TITLE    Direct Submission
JOURNAL  Submitted (16-MAR-1999) Shin-ichi Ayabe, Nihon University,
         Department of Applied Biological Science, Kameino 1866, Fujisawa,
         Kanagawa 252-8510, Japan (E-mail: ayabe@bri.nihon-u.ac.jp,
         Tel:81-466-84-3703, Fax:81-466-80-1141)
FEATURES
          Location/Qualifiers
          1..1500
            /organism="Lotus japonicus"

```

```

/mol_type="mRNA"
/db_xref="taxon:34305"
/clone="LjCYP-2"
/tissue_type="root"
/country="Japan: Gifu"
1..1500
/codon_start=1
/product="cytochrome P450"
/protein_id="BAA93634.1"
/db_xref="GI:7415996"
/translation="MDIISFLYSLSFYVALFAIKLFLGSRKFKMLPQPTSLPIIGN
LHLKRPURHTFRALSEKYGDFVSLWFGNRLVSVVVSFADVOECFTKNDVVLNAPRF
LSGYIFNYTTLSTSYGERHNLRLITSLVLSNHRINSFSPRRDVTLLIRLKA
EDSKNFSEVELTSRFDMTNMIRMSIKRYTGEDCDMTLQESAFRDMVTLLQ
LSGANNKADFWPILRLVDFEGLEKRVKIGSKTDFRFLGLQEHQDKQRTANTWIDH
LTLQESPEVYVTDQILKGLALAMLAGTDSASVLTWSMNCVNLNYPVLEKIKLELD
THVQDRLVDESDDIPKLYLKNVINEHLRLTPAPLLPHSASDDCTTGGYKVPDRTI
VLINAWLRDPQLWTEATTFKPERFDKGBLEKLIPEFLGRRRACFGELLALRAISM
LALLIQCFDWRVSDDEIDMGDRGFLVLMKSIPIVKAMCKSRPVNNVFK"
BASE COUNT      385 a      378 c      353 g      384 t
ORIGIN
Query Match      38.7%; Score 720.2; DB 8; Length 1500;
Best Local Similarity 68.8%; Pred. No. 3.3e-152;
Matches 1024; Conservative 0; Mismatches 453; Indels 12; Gaps 2;
QY      75 TCTCATAGCTGTCTCTTTCTCTGGTTCATCTCTCGCGTGAAGTTTGTTCCTCAAGCA 134
Db      20 TTTACTATTTCTCTTCTACGTAGTTTGTGGCATATCAACTCTTCTCTGGCTCAA 79
QY      135 GAAATTTGAGAAACATACCAAGGTCTCTCTCTCTTCCCATATATAGAGAAACCTTAACC 194
Db      80 GAAAGTTCAAAAACCTCTCTCCCGTCCAACTCACTTCCCATATTTGGCAACCTTCACC 139
QY      195 TCCTCGACACGCCAATCCACCGTTTCTTCCAAAGCATGTGCAACAGTACCGCAAGTGG 254
Db      140 ACCTCAAAAGCCCCCTCCACCGCACCTTCAGGGCGCTGTCTGAGAAGTATGCGCAGCT 199
QY      255 TTTCCCTCTGGTTCTGGTTCAGGTCTGGCGCTGTGTCTCTCTCTCCAAACAGCATACCAAG 314
Db      200 TCTCCCTATGTTTCGAAACCGTCTTGTGGTGGTGTCTCTCTCTCGCGATGTCAGG 259
QY      315 AATGCTTCAACAAACAGACGTTGCTTGGCCAAACCGGCTACCTTCTCTCTCGGAAAT 374
Db      260 AATGCTTCAACAAACAGACGTTGCTTGGCCAAACCGGCTACCTTCTCTCGGAAAT 319
QY      375 ACATCTCTTACAACAACACACCGTAGCTCTGTCTCCCAAGGAGTGGTGTGGCCAAAGT 434
Db      320 ACATCTCTTACAACAACACACCGTAGCTCTGTCTCCCAAGGAGTGGTGTGGCCAAAGT 379
QY      435 TCCGCGCATCACCGCCCTGGACGCTCTCTCCACGACGCGTCTCACCTCTCTCCGAA 494
Db      380 TCCGCGCATCACCTCTCTTGGATGTCTCTCCACACCGCATCAACAGCTTCTCCCCGA 439
QY      495 TCCGAGCGACGAGACGAAGCGTCTGATGACAGAGTTGGTGTGGCCAAAGTCTCGAACG 554
Db      440 TCCGCGCGATGAACCAACACGCGTGTATAGAGCTGGCTGAGGACTCTGTGAAGAAC- 498
QY      555 AGGAGAGATTTGGCGGAGTGGAGATTAGTTGATGTTCACGACTTAATTAACAACA 614
Db      499 -----TTCTCCGAAGTGGAGCTCACTCAAGGTTCTTCGACATGACCTTCAACA 550
QY      615 TAATGAGGATGATATCGGGGAAGAGGTTTTTACGAGAGGAGAGTGCAGATGAAGACGTTG 674
Db      551 TCATGAGGATGATATTCGCGGAAGAGGTTACTATGAGAGGATTTGACATGACCGAGCTTC 610
QY      675 AGGAGCGAGGAGGTTTCAAGAGACTGTGACAGAGAAATTTGGAACTCATGGCTTGGCTA 734
Db      611 AGGAGCGAGGAGGTTTCAAGGACATGTTGACCGAGTGTCTGAGTTGTCGCGGGGGAATA 670
QY      735 ACAAGGAGATCACTTGGCTTTCTCAGGTGGTTCGATTTTCAGATGTGAGAGAGCGCT 794
Db      671 ACAAGGCGGATTTTATGCTTATTTCTCAGGCTTGTGATTTTGAAGGGTGTGAGAGAGGG 730

```

Overkamp, S.		Direct Submission	
TITLE		Submitted (22-APR-1999) Overkamp S., Institute for Plant	
JOURNAL		Biochemistry and Biotechnology, University of Muenster,	
		Hindenburgplatz 55, Muenster, 48143, GERMANY	
COMMENT		Related sequence: AJ012581 possible allele.	
FEATURES		Location/Qualifiers	
source		1. 1781	
		/organism="Cicer arietinum"	
		/mol_type="mRNA"	
		/cultivar="ILC 3279"	
		/db_xref="taxon:3827"	
		/clone_lib="lambda Zap"	
		/note="heterotrophic cell suspension culture"	
gene		1. 1781	
		/gene="cyp81E3v2"	
CDS		66..1562	
		/gene="cyp81E3v2"	
		/BC_number="1.14.14.1"	
		/codon_start=1	
		/product="cytochrome P450 monoxygenase"	
		/protein_id="CAB41490.1"	
		/db_xref="GI:4688640"	
		/translation="MGIFSFAYSLFYLSIFIFRLLFQSRKFNLPDPPSPFIIGN FHLKRLPRLHRTFKGLSKTYGDIISLWFSRLVWVSSISEFQCCCTKNDVLANRPF LSGKTYFVNYTTLGSTYSSEHNRNLRITSLVLSNHRINNFSGVRDETRQLITKLA EDSSISPAVELSPRPDMTFNNIMRMISGKRYGEDCDMSDLOEASQFRDMVTELLQ LSGANKKTDFFMLLAFDFENLEKRVNIADKTDAPLRLGLLEQQRNKERTNTMIDLH LNLOSPQEPYTDQIIKGLAMLLAGTSSAVTLEWMSNLLNHPVLEKVRDELDT HYQDRLVDESPLKLYLKNVIYETLRCTPAPLLPHSTSDDCIIIGGYKVPDRDTIV LINAIAHRDPKSWSEATTFKPERFDKEIEKVIKAFGMGRVCPGEALALRTISMTL ALLIOCFDWKRTSDMIDMAERDGFVLTKLVLKAMCKTRPVVNKIFK"	
BASE COUNT		557 a 344 c 332 g 548 t	
ORIGIN		Query Match 38.5%; Score 716.6; DB 8; Length 1781; Best Local Similarity 67.5%; Pred. No. 2.2e-151; Matches 1026; Conservative 0; Mismatches 484; Indels 9; Gaps 1;	
Qy	48	CAAGCATGGGAATGTTGTTGGTGGTCTCATACGCTGCTCTTTCTTGGTTCATTATCC	107
Db	58	CAATCATATTAGGGGATCTTTTCTCTTCGCATATCTCTCTTTTATCTTCTATATTT	117
Qy	108	TCGGCGCTGAAGTTGTTTTCAAAGCAGAAATTTGAGAAACATACCAACAGGTCTCTC	167
Db	118	TCATTTTCAGGCTTTTGTTCCAATCAAGAAATTTCAAAAATCTCCACAGGTCCACCTT	177
Qy	168	CTCTTCCCAATAAGAAACCTTTAACTCTCTGAAACAGCAATTCACCGTTCTTCCAAAC	227
Db	178	CTCTTCTATAATTGGCAACTTCCCATCATCTCAAAAGCTCCCGCTCCACCGCACTTTCAAAG	237
Qy	228	GCATGTCGAAACAGTAGCGCAAGTGGTTCCTCTGCTGCTGCTGCTGCTGCTGCTG	287
Db	238	GACTTTCTAAACATATGGTGATATCATTTTCACTTTGGTTTGGATCACTGCTGCTG	297
Qy	288	TCATCTCTCTCCAAACAGCATACCAAGATGCTTCCAAACACAGCACTGCTTGGCCA	347
Db	298	TTGTTTCTCTCTCTCTGATTTCAACATGTTTACAAAACGATGTTGCTCTAGCTA	357
Qy	348	ACGGGCTACCTTCTCTCTCGGAAATATACATCTTCTACAAACACACCGTAGGCTCT	407
Db	358	ATAGGCCACCGGTTCTCTCTGGAATAATACATCTTCTACAACTATACGACTTTTAGGATCCA	417
Qy	408	GCTTCCACCGGAGCACTGGCGCAACTCGCGCATCACCGCTTCCCGCTGAGCTCTCTCA	467
Db	418	CATCTCTACAGTAGCAATTTGGCGCAACTCTGCTGATATCACTTCGCTGATGTTCTTTCAA	477
Qy	468	CGCAGCGCTCCACTCTCTCTCCGGAATCCGAGCGACGAGACGAGAGCGCTCTGATGAGA	527
Db	478	ACCACCGTATCAACAACCTTCTCTGGAGTCGAGAGGATGAGACTCAACGACTTAATACAA	537
Qy	528	GGTTGGTGTGGCCAAAGAACTCGAAACGAGAGAGTTTGGCGAGTGGAGATTAGTTCGA	587

327 TTTTGGCAATCGTATTCACTCGCGCAAAACCAAGTATCTTGGCTACAAACACAAACA 386
 399 TAGGCTCTGCTCCACGGCAGACACTGGCGCAACCTCCGCGCATCACCGCTCGACG 458
 387 TGGTTTCCATCGTACGGAGATCTTTGGCGCAACTTGGCGCTATAGTTCAATTGAAA 446
 459 TCCTCTCCACGACGCGCTCACCTCTTCCGGAATCCGGAGCGAGCAGCAGCAAGCGTC 518
 447 TACTCTCCACTCACCGTCTCACTCTTTTCCGGAATCCGGAAGACGACGATGAGGT 506
 519 TGATGACAGAGTTGGTGGCGCAAGAACTCGAACGAGGAAGATTTCGCGGAGTGAGA 578
 507 TGATCGGAAAA-----CTTGCTAAAAAATCTCAC---AAAGATTTCACAAAAGTTGAGC 557
 579 TTAGTTCGATGTTCAACGACTTAATCTTACACACATAATGAGGATGATATCGGGGAAGA 638
 558 TTAGTCCCTTTTGGTGAGCTAACATTTAACTATATGAGTGGTTTGTGGTAAC 617
 639 GGTTTTACGAGAGAGAGTGAGATGAAGAACGTTGAGGAAGCGAGGAGTTTCAGAGAGA 698
 618 GATTTTATGTTGAAGATCCGATGGAACAAATGCTGAGGAGGCAAGAAGTTTAGAGACA 677
 699 CTGTGACAGAAATGTTGAACTCATGCGCTTTGGCTTAACAGAGGAGATCATTGCTTTCC 758
 678 TGATGAATGAGATGACAGAGTTTGGGTTGGGCTCCAACTTGGTGATTTTGTGCTTTAT 737
 759 TCAGTGTGTTGATTTTTCAGAAATGGAAGAGCGTTAAAGAGTATCAGTAAGAGGTACG 818
 738 TCAAGTTGTTGATTTTGTAGTGGTGTATAGAAGTTTAAAGAGTTGGGAGGAAATGG 797
 819 ATTCATCTTGAATAAGATCCTTCATGAGAACCGTGCAGCAATCAACCGCA---GAATT 875
 798 ATGCATTTTCAAGGACTCGTTGACGAGCATCGTATACAGAGACAAAATTTGATA 857
 876 CCATGATCGATCATCTCTCAAACTGGAAGAGACCGAGCTCAGTACTACACTGACCAAA 935
 858 CCATGATGATCATCTACTTTTCAATGCAAGAATCACAACCTGATTATTATAGTGACCAAA 917
 936 TCATCAAAAGGCTTGCTCTGCGCATGCTTTTGTGGAACTGACTCATCACTGGGACTT 995
 918 TTATCAAGGGCTTATATGAGCATGATGATGGCTGGAAAGAACTTCATCTGTGACAT 977
 996 TAGAGTGTCTAATCTAATTTATGATATCCAGAGAGTGTGGAAGAGGCAAGAGATG 1055
 978 TAGAGTGGCGCATGCTAATTTATTAACCATCCAGAAATTTTGGAAAAGCAAGATTG 1037
 1056 AATTGGACACTCAAGTGGGACAGACCGCTTGTAAATGAGTCAAGCTTCCAAAACCTTC 1115
 1038 AATTGGACAATCACATTTGCAAGAGACCTTAATTTGAGGAAGCAGCAAGCTTACCAATTTAC 1097
 1116 CATATCTTAGGAAGATCATCTTGAGACACTTAGGTTGTACCCCGCGCCCAATTTCTAA 1175
 1098 AATACTTTCAAAACATTATCTCGAGACTTTTGGGTTGCATCCACCGGTACATTTGT 1157
 1176 TACCTCATGTCTTTCAGAGATATTAACAATGAAGGATTCAATATCCACGAGACAAA 1235
 1158 TGCCACATTTATCATACACAGATTGATATAGGAGGCTATGATGTGCAAGTAAACCA 1217
 1236 TTGTGATCATTAATGTTGGGCGATGACAGAGATCTCTCAGTTGTGGAATGATGCCACAT 1295
 1218 TGTGTATGGTGAATGATGGGCTATTCAATGGGATCCAAATTTTGGGGCTGACCCAATGA 1277
 1296 GCTTTAAACCTCAGAGGTTTG---ATGTGGAAGAGAGGAGGAAAAAGTTGTGTAGCATTTG 1352
 1278 GCTTTAAGCCCGAGAGGTTTGAAGATGTTCAAGGGGACATTGTTGGCTTTATCTCTTTG 1337
 1353 GCATGGGAAGAGGCTTCCACAGAGAACCCATGGCTATGCAAGTGTTCAGCTTTACTT 1412
 1338 GGAATGGGAGGCGGCTTGTCTCTGGTCCGGCTAGCGCTTAAGAACAATTGGGCTTAACTT 1397
 1413 TGGGATTTGTTCAATGTTTGTGCTGGAACGAGTGAAGTGAAGGAAAGCTTGAATGA 1472

Db 1398 TGGGCTGTTGATTCAATGCTTTAGTGGAAAAAGGANTGGGAAGGAAGTGTGACATGA 1457
 QY 1473 CAGAGAACAAATGGATCACCTTGTCAAGGTTAATTCATTGGAGGCCATGTGCAAGGCTC 1532
 Db 1458 GTGAGGAAGCGGAACACTGTGGTGGCCAGGCCATTTCCATTAGAGGCCCATGCAAGAGCC 1517
 QY 1533 GCCACTTCCCACTAAATTTGGAATTTAA 1561
 Db 1518 GACCAATAATCAACAGATTTTCAATTA 1546
 RESULT 7
 AP006142/c 104321 bp DNA linear PLN 19-MAR-2003
 LOCUS Lotus japonicus genomic DNA, chromosome 2, clone:LJT08A24, TM0250,
 DEFINITION complete sequence.
 ACCESSION AP006142
 VERSION AP006142.1 GI:29122779
 SOURCE HTG.
 ORGANISM Lotus japonicus
 LOTUS japonicus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteeae;
 Lotus.
 REFERENCE 1
 AUTHORS Kaneko, T., Asamizu, E., Kato, T., Sato, S., Nakamura, Y. and Tabata, S.
 TITLE Structural Analysis of a Lotus japonicus Genome. III. Sequence
 Features and Mapping of Sixty-two TAC Clones Which Cover the 6.7 Mb
 Regions of the Genome
 JOURNAL DNA Res. 10, 27-33 (2003)
 REFERENCE 2 (bases 1 to 104321)
 AUTHORS Sato, S.
 TITLE Direct Submission
 JOURNAL Submitted (12-DEC-2002) Shusei Sato, Kazusa DNA Research Institute,
 Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu,
 Chiba 292-0818, Japan (E-mail: ssato@kazusa.or.jp,
 URL: http://www.kazusa.or.jp/, Tel: 81-438-52-3935 (ex. 2337),
 Fax: 81-438-52-3934)
 FEATURES
 1. 104321
 Location/Qualifiers
 /organism="Lotus japonicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:34305"
 /chromosome="2"
 /clone="LJT08A24"
 /clone_lib="LJT library"
 /note="TAC clone: TM0250"
 BASE COUNT 31736 a 18441 c 19683 g 34461 t
 ORIGIN
 Query Match 22.4%; Score 415.8; DB 8; Length 104321;
 Best Local Similarity 57.3%; Pred. No. 5.5e-83;
 Matches 896; Conservative 0; Mismatches 567; Indels 100; Gaps 4;
 QY 78 CATAAGCTGTCTTTTCTGTTCTATTCCTGGCGTGAAGTTGTTTTCGAAGAGAA 137
 Db 8534 CCTACATCTCTGTTCTTCAITTTTCTCATCACAGCCAAGTTCTTACTCTCCAGAGAA 8475
 QY 138 AATTGAGAACATACCACAGGTCCTCTCTCTCCCATATAGGAAACCTTAACTCC 197
 Db 8474 GATTCAAGAACCTTCCACAGGCCACGCCCTCTTCCCTTAATCGGCAACTCCACAGC 8415
 QY 198 TCGAACAGCAATCCACCGTTCTTCCAAACGCGATGTGAAACAGTAGTCGGCAAGTGTGTTT 257
 Db 8414 TAAACCAACGATTCACACTTTTCCAGCCCTATCTCAGAAATATGTTCCCATTTTCT 8355
 QY 258 CCCTCTGTTGCGTTTCACTGTCGGCGTTGTATCTCTCTTCCAAAGAGATACCAAGAT 317
 Db 8354 CTTTCCGTTGCGGTTCCCTACCGGTGCTGCTTTCTCTCTGCTCCGCGCGGAAGAT 8295
 QY 318 GCTTTCACCAACACGAGAGTGTGCTTGGCCACCGGCTACCTTCTCTCTCGGGAATAACA 377

Db 8294 GCTTACCAAAAGCAGCATCATCTTCCGCAACCGTGTCCACTCTCTCAAGACCAAGCACC 8235
 QY 378 TCTTTACAAACACACACCTAGGCTCTCTGCTCCACGGCGAGCACTGGCGCAACCTCC 437
 Db 8234 TCGCTACGACAGCAGCAGCAGCTCATCACCGCTTCTTACGGCGACCACTGGCGCAACCTCC 8175
 QY 438 GCGCATCACCGCCCTGGAGTCTCTTCCACGAGCGCGTCCACTCTTCTCGGAATCC 497
 Db 8174 GCGGTATCAGCTCAATTTGAGATCTCTCAACCCACCGCCTCAACTCTCTTCTCGGAGTCC 8115
 QY 498 GGAGCGACGAGACGAAGCGTCTGATGACAGAGTGTGTGCTGGCCAGAGACTCTCAACGAGG 557
 Db 8114 GAAAGGACGAGACCATGAGGCTGCTCGGAGCTCGCTC-----ATPACAAAG 8067
 QY 558 AAGAGTTTGGCGAGTGGAGATGATTGCTGATGTTCAACGACTTAACTTACAAACAAATAA 617
 Db 8066 ATGAGTTTCGGAAGATTGAGTCCAGTCCAGTCTGTTGCTGAACCTAACTTCAACACCATCA 8007
 QY 618 TGAGGATGATTCGGGAAGAGTTTACGAGGAGAGTGTGATGAGTGAAGAGTTGAGG 677
 Db 8006 TGAGATGGTGTGTGGCAACCGTACTACCGCGAGGAGTTTGACGGAACCGAGCGCGAGG 7947
 QY 678 AAGCGAGGAGTTTCAGAGAGACTGTGACAGAAATGTTGGAACCTCATGGGCTTGGCTAAACA 737
 Db 7946 AGCGAAGAAATTCAGAGATGTCTATGATGAGATGCAACGTTTGGATTAGGCTTGGTC 7887
 QY 738 AGGAGATCACTTGCCTTTCTCAGGPGTTCGATTTTCAGATGAGAGAGCGCTTAA 797
 Db 7886 TCGGTGAATTCGTGCTCTTTTTCAGGTGTTGTTGATTTTCAGCGGAGAACCCAGCGGATTAC 7827
 QY 798 AGAGTATCAGTAGAGTACGATTCATCTTGAATGAATGATCCCTTCATGAGAACCGTGCCA 857
 Db 7826 AGAACTTGGGGGAGAAATGGAGCGACTCTTCAAGGAGCTCATCGATGAACCTCGGAATA 7767
 QY 858 GCAATGACCCCGCAG--AAATTCATGATCATCATCTCTCTCAAACTGCAAGACAGCCAGC 914
 Db 7766 TGGGGGACCAAAATCAACACGATGATGATCATCTACTTCTTACAGAGACACAGC 7707
 QY 915 CTGAGTACTACATGACCAATCAATCAAGGCTTGTCTGCG-----956
 Db 7706 CTGAGTACTACACTGACCAAAATATCAAGGGCTTATTATGGTGGTGGGAAATGTGAT 7647
 QY 957 -----956
 Db 7646 ATGTCAATAAATTGATTAATAATAAATAATTTATATTTATTAATGTTGTTGTTTGG 7587
 QY 957 ----CCATGCTTTTGTGTGAACTGACTCATCACTGGGACTTTAGAGTGGTCAATATCT 1012
 Db 7586 CAGTCTGTATAGTGGCGGACAGAACTTCAGTACACATAGATGGGTGATGGCG 7527
 QY 1013 AATTATTTGAATCACCCAGAGGTGTTGAAGAGGCAAGAGATGAATTTGGACACTCAAGTG 1072
 Db 7526 GAGCTGTGAACCATCCGCGGTTGTTGGAGAAAGCGAGGAATGAGTTGGAGTCACAAAT 7467
 QY 1073 GGCACAGACCGTGTAAATGAGTCAGACCTTCCAAACCTTCATATCTTAGAGAGATC 1132
 Db 7466 GGAATGAGCGGTAATGGAGGAGCGGAGGTAACTAAAGTGCCTTACTTTGGAGAAATA 7407
 QY 1133 ATCTTTGAGACACTTAGTGTGTACCCCGCGCCCAATTTCTAATACCTCATGTGTCTTCA 1192
 Db 7406 ATCTGTGAGACTTTCGCTGTGATCCAGAGTTCGAGTGTGTGTACCAATTAATCTCTCT 7347
 QY 1193 GAAGATATTAATTTGAAGATTAATATATATCCACGAGACAAATGTTGATCATTAATGTT 1252
 Db 7346 GAGGATTTGACGGTGGTGGTATCAGTCCGCGTAAATCAATGTTGATGTTGATGGTGAATGCG 7287
 QY 1253 TGGGGATGAGAGAGATCTCTAGTGTGGAATGATGCCACATGCTTTAACTTGAAGG 1312
 Db 7286 TGGGCTATACAGAGGAGCCCAATTTGTGGTGTGATCCGTTGCAATTTGTCCTGAGAGG 7227
 QY 1313 TTGTG---ATGTGGAAGGAGAGGAGAAAGTGTGTAGCATTTTGGCATGGGAGAGCGGT 1369
 Db 7226 TTGTAGATGTTGTTGGTATGCGGCTGTGGTCAATCTTCTTGGATGGGAGAGGTCA 7167

QY 1370 TCCCCAGGAGAACCCATGCTATGCAAAAGTGTGACGTTTACTTTGGGATTTGTTGATCAA 1429
 Db 7166 TGTCTCTGGTTCGGGCTTGGCCAGAGAACTATATGGTCTGACTCTGGGCTCGTTGATTCAG 7107
 QY 1430 TGTTTTGTACTGAAACGAGTAAAGTGAAGAAAGCTTGTATGATGACAGAGAACTTTGGATC 1489
 Db 7106 TGTCTTGTGATGGAAGAGATTGGCAGCGGAAGAGATTGACATGATGATGAAAGACGTGGAAT 7047
 QY 1490 ACCTTGTCAAGGTTAATTCCTATTGAGGCGCATGTGCAAGGCTCGCCACTTGCACATAA 1549
 Db 7046 ATCATACCTTAAGCTATTCCTTTGAGGCGCTTTGCAAGCTCTCTCAATCATCAACAG 6987
 QY 1550 ATT 1552
 Db 6986 ATT 6984

RESULT 8
 BT002874
 LOCUS
 DEFINITION
 Arabidopsis thaliana clone RAFL15-01-B10 (R20394) putative cytochrome p450 family protein (At4g37320) mRNA, complete cds.
 ACCESSION
 BT002874
 VERSION
 BT002874.1
 GI:27754486
 FLI CDNA.
 SOURCE
 Arabidopsis thaliana (thale cress)
 ORGANISM
 Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1693)
 YAMADA,K., CHAN,M.M., CHANG,C.H., DALE,J.M., HUAN,V.W., LEE,J.M., ONODERA,C.S., QUACH,H.L., TANG,C.C., TORIUMI,M., WONG,C., WU,H.C., YU,G., YUAN,S., CARNINCI,P., CHEN,H., CHEUK,R., HAYASHIZAKI,Y., ISHIDA,J., JONES,T., KAMIYA,A., KAWAI,J., KIM,C.J., NARUSAKA,M., NGUYEN,M., PALM,C.J., SAKURAI,T., SATOU,M., SEKI,M., SHINN,P., SOUTHWICK,A., TRIPP,M.G., WU,T., SHINOZAKI,K., DAVIS,R.W., ECKER,J.R. and Theologis,A.
 Arabidopsis Full Length cDNA Clones

TITLE
 JOURNAL
 Arabidopsis Full Length cDNA Clones
 REFERENCE
 2 (bases 1 to 1693)
 YAMADA,K., CHAN,M.M., CHANG,C.H., DALE,J.M., HUAN,V.W., LEE,J.M., ONODERA,C.S., QUACH,H.L., TANG,C.C., TORIUMI,M., WONG,C., WU,H.C., YU,G., YUAN,S., CARNINCI,P., CHEN,H., CHEUK,R., HAYASHIZAKI,Y., ISHIDA,J., JONES,T., KAMIYA,A., KAWAI,J., KIM,C.J., NARUSAKA,M., NGUYEN,M., PALM,C.J., SAKURAI,T., SATOU,M., SEKI,M., SHINN,P., SOUTHWICK,A., TRIPP,M.G., WU,T., SHINOZAKI,K., DAVIS,R.W., ECKER,J.R. and Theologis,A.
 Arabidopsis Full Length cDNA Clones

TITLE
 JOURNAL
 Direct Submission
 Submitted (15-JAN-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-Length cDNA': Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.)

COMMENT
 The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Huan,V.W., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

YAMADA,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.
 Location/Qualifiers


```

source      1. .1693
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosomes="4"
/clones="RAFL15-01-B10 (R20394)"
/notes="This clone is in a modified pBluescript vector2
(lambda PS) as a BamHI/XhoI insert.
ecotype: Columbia"
1. .1693
/gene="At4g37320"
5'UTR      1. .27
/gene="At4g37320"
misc_difference 1. .21
/gene="At4g37320"
/notes="not present in genomic sequence"
28. .1515
/gene="At4g37320"
CDS        /codon_start=1
/evidence=experimental
/product="putative cytochrome p450 family protein"
/protein_id="AA022691.1"
/db_xref="GI:27754487"
/translation="MDLTQILLISFLPTISIKLLTKSNRKNLPSPNYPPLVPIGH
LHLKQVHRTHFSISKSLGNAPIFHLRLGNRLVYVISHSIAEBCTKNDVVLNRP
DIIMAKHVGNFTNMAASYGDHWRNLRIAAVEIFSSHRISTFIRKDIIRLLIHA
LSRSLHGFVEVLEKSLTINLAFNNIMVAGKRYGTGTEDNDEAKLVRELIETHA
GAGSGLNADLYLSINWYFNENOTKILGNLDRLVLOKLVDEKRAEKKGOTLLDHLS
FOETPEVYDVLPIKGIILALVLAGTDTSSVTLWAMSNLHNPHEILEKARABIDKI
GSDRLVESIDVNLHLYQNIIVSETRLIYPAVPLLPFSSDECKVAGYDMPRTLILT
NVAMHRDPGLWEEPEFKPEKEGEARKLMPFGMRRCAPGAEILKRLVSLALGC
LIQSFWEYRVAELVDNTEGEGITMPKATPLRACAKAIVGKTI"
1516. .1693
/gene="At4g37320"
3'UTR      1678
misc_difference 1678
/gene="At4g37320"
/notes="not present in genomic sequence"
BASE COUNT 482 a 394 c 374 g 443 t
ORIGIN

Query Match      21.2%; Score 394.8; DB 8; Length 1693;
Best Local Similarity 56.9%; Pred. No. 1.6e-78;
Matches 817; Conservative 0; Mismatches 592; Indels 27; Gaps 4;

Qy 162 CTCCTCTCTTCCCATATAGGAAACCTTAACTCTCTCGAACACGCCCAATCCACGGTTTCT 221
Db 134 CTTACCCCTTTACCGGTTATCGGCACCTCCACCTCTCAAGCAACCGGTCCACCGACAT 193
Qy 222 TCCAACGGATGCGAAACAGTACGGCAACG-----TGGTTTCCCTCTGGTTCCGTTAC 275
Db 194 TCCACTCCATCTCTAAGTCCCTTGGAAATGCTCCAATCTTCCACCTCCGCTGGGAAACC 253
Qy 276 GTCTGGCGGTTGTCTATCTCTCCAAAGCATACCAAGATGCTTCACCAACACGACG 335
Db 254 GCCTCGTTTACGTATCTCTCACTCACTCATAGCGGAAGAAGTTTACCAAAAACGACG 313
Qy 336 TTGCCTTGGCCAAACGGGTACTCTCTCTCGGGAATAATACATCTTCTACAAACACCA 395
Db 314 TCGTTCTTGGAAACCGTCTCTGATATCATCATGCGCCAAACACGTCGGCTATAACTTTACCA 373
Qy 396 CCGTAGGCTCTGTCTCCACGGCGAGCACTGGCGCACTCGCGGCATACCGCCCTGG 455
Db 374 ATATGATTGCGAGCATCTTACGGCGACCACTGGAGGAATCTCGCGGCATCGCCCGCTCG 433
Qy 456 ACGTCTCTCCAGCGAGCGGTCCACTCTCTTCGGGAATCCGGAGCGACGACGAGCAAGC 515
Db 434 AGATATTTCTTCTCATAGATCAGTACCTTTCTTCTTATCCGTAAAGACGAGATCCGCAC 493
Qy 516 GTCTGATGACAGAGTTGGTGTGGCCCAAGAACTCGAACGAGGAAGAGTTTTCGGCGAGTGG 575
Db 494 GGCTCATACACATCTCTTAGAGAC-----TCCCTTGCACGAGTTTGTAGTAGTGG 544
Qy 576 AGATTAGTTCGATGTTCAACGACTTAACCTTACAAACATTAATGAGGATGATATCGGGGA 635

```

RESULT 9
BT004449
LOCUS

Arabidopsis thaliana clone U20394 putative cytochrome p450 family
protein (At4g37320) mRNA, complete cds.

1519 bp mRNA linear PLN 14-FEB-2003

545 AGTTCGAATCGTTATTAAACCACTTGGCATTCACCAACATCATCATGATGTTAGCGGAA 604
Qy 636 AGAGGTTTTACGGAGAGGAGTGAAGAACGTTTGAGGAACCGGAGGAGTTTCAGAG 695
Db 605 AACGATATTAACG-----TACCGGTACAGAAGACAATGATGAAGCCAGCTCGTAGGG 658
Qy 696 AGACTGTGACAGAAATGTTGGAACCTCATGCGGTCTGGCTTAACAGGGAAGATCACTTGCCTT 755
Db 659 AACTTTATAGCGAGATAAATGCGCGCGCGGTTCTTGGAAATTTTGGCTGATTATCTTCCGT 718
Qy 756 TCCTCAGGTGGTTCGATTTTCAGATGTGGAGAACGCTTAAAGAGTATCAGTAAGAGGT 815
Db 719 CCATAAATTTGGGTC-----ACAAACTTTGAGAACACAGACCAAAATCTTGGGAATCGAC 772
Qy 816 ACGATTTCATCTTGAATGAATCCTTTCATGAGAACCGTCCAGCAATGACCCCGAGAATT 875
Db 773 TCGATAGAGTCTCTGCAAAAACCTGGTTGATGAGAAACGTGCAGAGAAAGAAAGGTCAAA 832
Qy 876 CCATGATGATCATCTCTCTCAAACTGCAAGAGACCCAGCTCAGTACTACATGACCAAA 935
Db 833 CTTTGATCGATCATTTGCTTTCTTCCAGAAACCGAACCCCGAGTACTATATCTGATGTCA 892
Qy 936 TCATCAAAAGCCCTTCTCTGCGCATGCTTTTGGTGGAACTGACTCATCACTGAGACTT 995
Db 893 TCATCAAAAGGAATCATCTGCGCTTTGGTTCTTTCGGGGACAGATACATCTTCAGTGACGT 952
Qy 996 TAGAGTGGTCATATCTAATTTATTTGAATCACCCAGAGGTGTTGAAGAAAGCAAGAGATG 1055
Db 953 TGGAAATGGCAATGTCAAAATTTGCTGAACCATCCAGAGATATCTTGAGAAAGCGAGACG 1012
Qy 1056 AATTGGACACTCAAGTGGGCAAGACCGCTGTTTAAATAGTTCAGACTTCCAAAACCTTC 1115
Db 1013 AGATCGATGATAAATCGGTTTCAGACCCGTTAGTTAGTGAAGATCAGATATTTGAAATCTCC 1072
Qy 1116 CATATCTTAGGAAGATCATCTTTCAGACACTTAAAGTTGTACCCCGCGGCCCAATTTCTAA 1175
Db 1073 ATTATCTCCAAACCATTTGTTCAGAAACATTAAGTTGTATCTGCGGTTCCACTACTAC 1132
Qy 1176 TACCTCATGTCTTCAGAAAGATATTAACAATTTGAAGGATTCATATCCACGAGACACAA 1235
Db 1133 TCCCTCATTTCTCATCGGATGAATGTAAAGTGGCGGCTACGATATGCCACGTCGCACGT 1192
Qy 1236 TTGTGATCATTAATGTTGGGCAATGACAGAGATCTCAGTTGTGGAATGATGCCACAT 1295
Db 1193 TGTATTAAACAAACGATATGGCGGATGATAGATCCAGTTTATGGGAAGACCCAGAGA 1252
Qy 1296 GCTTTAAACCTGAGAGGTTTGTATGTAAGAGGAGAGAGAAAGTTGGTAGCATTTGGCA 1355
Db 1253 GGTTCAGCGGAGAGGTTTCGAGAAAGAGAGAGGCTCGAAAGCTAATGCCGTTTGGGA 1312
Qy 1356 TGGGAAGAGAGGCTTCCCGAGAGAACCCATGGCTATGCATAAGTGTGAGCTTTTACCTTTGG 1415
Db 1313 TGGGACGACGAGCTTGTCTGGAGCTGAGCTTGGGAAGCGGTTAGTGAGCCTTGTCTCTTG 1372
Qy 1416 GATTGTTGATTCATGTTTGTGTAAGAGAGTAAAGTGAAGAAAGCTTGTATATGACAG 1475
Db 1373 GGTGCTGATTCAGTCTTTTCGAGTGGGAGAGATTTGGTGCAGAACTTTGTGACATGACTG 1432
Qy 1476 AGAAACAATTTGGATCAGCTTGTCAAGGTTAATTCATTGGAGGCGCATGTCAAGGCTCGCC 1535
Db 1433 AAGCGAAGGATTCATATGCTTAAAGCTACTCCGTTGGAGCTATGTGCAAGGACGTG 1492
Qy 1536 CACTTCCCACTTAAATTTGGAATTTAATTAATAGTATTTTATTTTATTTTGGTAAACTT 1591
Db 1493 CCATTGTTGTTAAACGATATAATTTACGCTATGTAGTACTATTTGTGTAGAACCT 1548


```

Db      986 AGATCGATGATAAAATCGGTTTCAGACCGGTTAGTTGAAGAATCAGATATTTGTAATCTCC 1045
Qy      1116 CATATCTTAGAGAGATCATCTTGAGACACTTAGGTTGTACCCCGCCGCCCAATTTCTAA 1175
Db      1046 ATTAATCCAAACAATTTGTGAGAAATATAGTTTGTATCTCTGCGGTTCCACTACTAC 1105
Qy      1176 TACCTCATGTGCTTTCAGAAAGATATTAACAATTGAAGGATTTCAATATCCACGAGACACAA 1235
Db      1106 TCCCTCATTTCTCATCGATGAATGTAAAGTGGCGGCTACGATATGSCACAGTCGACGTT 1165
Qy      1236 TTGTGATCATTAATGTTGGGCGATGCAGAGAGATCTCTAGTTGTGGAATGATGCCACAT 1295
Db      1166 TGTATTAAACAAAGTATGGCGGATGCATAGAGATCCAGGTTTATGGGAAGAGCCAGAGA 1225
Qy      1296 GCTTTAAACCTGAGAGGTTTGATGTGAAGAGAGAGAGAAAGTTGGTAGCATTTGGCA 1355
Db      1226 GGTTCACGCGAGAGGTTTCAGAAAGAGAGAGGCTCGAAAGCTAATGCCGTTTGGGA 1285
Qy      1356 TGGGAAGAGGCTTGGCCAGGAGAACCCATGGCTATGCAAGTGTGAGCTTTACTTTGG 1415
Db      1286 TGGGACGACGAGCTTGTCTGGAGCTGAGCTTGGGAAGCGTTAGTGAGCCTTGCTCTG 1345
Qy      1416 GATTGTTGATTCAAATGTTTGTGACTGGAAACGAGTAAGTGAAGGAGAAAGCTTGATGACAG 1475
Db      1346 GGTGCTTTGATTCAGTCTTCGAGTGGGAGAGAGTTGGTGCAGAACTTTGTGCACATGACTG 1405
Qy      1476 AGAACAAATTCGATCCTGCTCAAGGTTAATTAATTAATAGTATTTATTTATTTGGTAAAC 1589
Db      1406 AAGCGAAGGATCAGTATGCTTAAGCTACTCGTTGCGAGCTATGTGCAAGGACGCTG 1465
Qy      1536 CACTTGGCCACTAAATTTGGAATTTAATTAATTAATAGTATTTATTTATTTGGTAAAC 1589
Db      1466 CCAATTGTTGTAACAGATATAATTGACGTATGTAGGATCTATTGTGTAGAAAC 1519

RESULT 10
AX506493 1488 bp DNA linear PAT 27-SEP-2002
LOCUS Sequence 1188 from Patent WO0216655.
DEFINITION AX506493
ACCESSION AX506493
VERSION AX506493.1 GI:23387730
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
REFERENCE 1
AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
TITLE Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
JOURNAL Patent: WO 0216655-A 1188 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
source Location/Qualifiers
1..1488
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 408 a 360 c 336 g 384 t
ORIGIN
Query Match 21.2%; Score 393.8; DB 6; Length 1488;
Best Local Similarity 57.4%; Pred. No. 2.6e-78;
Matches 801; Conservative 0; Mismatches 567; Indels 27; Gaps 4;
Qy 162 CTCCTCCTCTTCCCATATAGAAACCTTTAACTCTCTCGAAGACCCCAATCCACCGTTTCT 221
Db 107 CTTACCCCTTTTACCGGTTATCGGCGACCTCCACCTCTCCACCAACCCGCTCCACCGAAT 166
Qy 222 TCCAACGATGTCGAACAGTACGGCAACG-----TGGTTCCCTCTGTTTGGTTTCAC 275

```

```

Db      167 TCCACTCCATCTCTAAGTCCCTTGGAAATGCTCCAATCTTCCACCTCCGCTTGGAAACC 226
Qy      276 GTCTGGCGGTTGTCATCTCTCTCTCAACAGAGATACAAAGAATGCTTCCACCAACACGAGC 335
Db      227 GCCTGTTTACGTCATCTCTCTCACTCCATAGCCGAGAGATGTTTCCACCAAAACGAGC 286
Qy      336 TTGCTTGGCCAAACCGGTACCTCTCTCTCGGGAATAATACATCTTCTACAAACAACACA 395
Db      287 TCGTCTTTCGGAACCGCTCTGATATCATATGCGCAACACATGCGCTGCTATTAATCTTACCA 346
Qy      396 CGGTAGGCTCTGCTCCACGCGAGCACTGGCGCAACCTCCGCGCATCACGCGCTTGG 455
Db      347 ATATGATTGACAGCATCTTACGCGACCACTGGAGGAATCTCCGCGCATTCGCGCGCTG 406
Qy      456 AGCTCTCTCCACGCGAGCGCTCCACTCTCTCCGGAATCCGAGCGACAGAGAAAGC 515
Db      407 AGATATTTCTCTCATAGATCAGTACCTTTTCTCTATCCGTAAGAGAGATCCGAC 466
Qy      516 GTCTGATGACAGAGTTGGTGGCCAAAGAACTCGAAAGAGAGAGTTTGGCGGAGTGG 575
Db      467 GGCTCATAAACATCTCTCTAGAGACT-----CCTTGACCGAATTTGTTGAAGTGG 517
Qy      576 AGATTAGTTCGATGTTCAACGACTTAACTTACAAACAATATAGGATGATATCGGGCA 635
Db      518 AGTTGAAATCGTTATTAACCAACTTGGCAATCAACAACATCATCATGATGGTAGCCGAA 577
Qy      636 AGAGGTTTTACGAGAGGAGAGTGAAGAAAGTTGAGGAAGCGAGGAGTTTCAGAG 695
Db      578 AACGATATTAACG-----TACCGGTACAGAGACATGATGAGCCAGCTCGTAGGG 631
Qy      696 AGACTGTGACAGAAATGTTGGAACTCATGGGCTTGGCTTAAACAAGGAGAGATCATCTTGCCTT 755
Db      632 AACTTATTAGCGAGATTAATGGCGCGCGCTTCTGGAAATTTGGCTGATATCTTCCGT 691
Qy      756 TCTCAGGTGGTTCGATTTTCAAGATGAGAGAGCGCTTAAAGAGTATCAGTTAGAGGT 815
Db      692 CCAATAAATTTGGTCT-----ACAAACTTTGAGAACCAAGACCAAAATCTTTGGGGAATTCGAC 745
Qy      816 AGGATTCCATCTTGAATTAAGATCTTTCATGAGAACCGTGCAGCAATGACCGCAGAAAT 875
Db      746 TCGATAGAGTCTCTGCATAAACTGGTTGATGAGAAACGTCGACAGAAAGAAAGGGTCAAA 805
Qy      876 CCAATGATCATCATCTCTCAAACTGCAAGAGACCCAGCTCAGTACTACTACATGACCAAA 935
Db      806 CTTTGATCATCATTTGCTTTCTTCCAGAAACCGAACCCGAGTACTACTACTGATGTCA 865
Qy      936 TCATCAAAAGGCTTCTCTGCGCATGCTTTTGGTGGAACTGACTCATCACTCACTGGGACTT 995
Db      866 TCATCAAAAGGATCATCTACTGCTTTGGTCTTTTGGTCTTTTGGGGGACAGATACATCTTCACTAGTACCGT 925
Qy      996 TAGAGTGGTCAATATCTAATTTTATTAATCAACCCAGAGGTGTTGAAGAAGGCAAGAGATG 1055
Db      926 TGGAAATGGCAATGTCAAAATTTGCTGAACCATCCAGAGATACTTGAGAAAGCGAGAGCGG 985
Qy      1056 AATTGACACTCAAGTGGGACAAAGCCGCTTGTAAATGAGTCAGACCTTCCAAACTTC 1115
Db      986 AGATCGATGATAAATTCGGTTCAGACCCGTTAGTTGAAGAATCAGATATTTGTAATCTCC 1045
Qy      1116 CATATCTTAGGAAGATCATCTTTGAGACACTTAGTGGTTGTACCCCGCGGCCCAATTTCTAA 1175
Db      1046 ATTATCTCCAAACATTTGTGTGAGAAACATTTAGTTTGTATCTCTGCGGTTCCACTACTAC 1105
Qy      1176 TACCTCATGTCTTTCAGAAAGATATTAACAATTAAGGATTTCAATATCCACGAGACACAA 1235
Db      1106 TCCCTCATTTCTCATCGGATGAATGTAAAGTGGCGGCTACGATATGCCACGTCGACAGT 1165
Qy      1236 TTGTGATCATTAATGTTGGGCGATGCAGAGAGATCTCAGTTGTGGAATGATGCCACAT 1295
Db      1166 TGTATTAAACAAACGATATGGGCGATGCATAGAGATCCAGGTTTATGGGAAGAGCCAGAGA 1225
Qy      1296 GCTTTAAACCTGAGAGGTTTGTGTGGAAGGAGGAGAAAGTTGGTAGCATTTGGCA 1355
Db      1226 GGTTCACGCGGAGAGGTTTCGAGAAAGAGAGAGGCTCGAAAGCTAATGCCGTTTGGGA 1285

```


497 TAAATTTGTCTAG-----CGAAAACTCTCAACATGGGTTTGTGAAGTGGAGATGA 547
582 GTTCGATGTTCAACGACTTAACATCAACACATAATGAGATCATATCGGGAAGAGGT 641
548 AACCATTTGTTATGTTTAACTATCAACACATCATTTAGNAATGGTGGCGGAANAAGAT 607
642 TTTACGGAGGAGAGTGCAGATGAAGAACGTTGAGGAGCGGAGGAGTTTCAGAGAGACTG 701
608 TCTACGGTGTGAACAG-----AGAACGACACAGGAGGTAAACACGTCGAGACAGCTAA 661
702 TGACAGAAATGTTGGAATCTATGGCTTGGCTTAACAAGGGAGATCATCTTGCCTTTCTCA 761
662 TTGCGGAGGTAGTTGTTAGCGCGCGCGAGGAGAAACGACGACAGACTACTCTCCATCTTC 721
762 GGTGTTTCGATTTTCAGAACTGTCAGAGAGCGCTTAAAGAGTATCAGTAAGAGGTACCAT 821
722 GTT-----ACGTAACGAACTATGAGAAACATGTTAAGAGTTAGCAGGTGCTGTTGATG 775
822 CCATCTTTGAATPAAGATCCTTCATGAGAACCGTGCAGCAATGACCGCAGAAATTCATGA 881
776 AGTTTTTGCAAAGTCTAGTTAAACGAGAAACGTTGGAGAAAGTAAAGGTAACACAAATGA 835
882 TCGATCATCTCTCAACTCAAGAGACCGACCTCAGTACTACATGACCAATATCATCA 941
836 TCGATCATTTTACTTCTCTACAAGAAACTCAGCCCTGATTACTACACGGATGTGATCATCA 895
942 AAGGCTTGTCTCGCCCATGCTTTTGTGTGAACCTGATCATCACTCAACTGGGACTTTAGAGT 1001
896 AAGGAATCATACTTGTATTGATACTTTCGCGGACTGATACATCAGCTTGGAAACGTTAGAT 955
1002 GGTCAATTAATTAATTTGAATCACCAGAGAGGTGTTGAAGAGCGCAAGAGATGAATTTGG 1061
956 GGGCGATGCGAATTTGTTGAACCATCCAGAAAGTATTACGAAAGCGCAAGACCGAATCG 1015
1062 ACATCAAGTGGGACAGACCGCTTGTAAATGAGTCAGACTTCCAAAACCTTCATATC 1121
1016 ACGACCAATCGGTGTAGACCGGTTAGTAGAGGACACAGACATTTGTGAACATTCCTTACC 1075
1122 TTAGAAGATCATCTCTGAGACACTTAGTTGTACCCCGCCCGCCCAATTCATATACCTC 1181
1076 TTCAGCATTTGTCTCAGAGACTTTACGGCTTTATCCGGTGGCTCCCAATGCTTTTACCCC 1135
1182 ATGTGCTTCAGAGATATTCAATTTGAAGGATCAATATCCACGAGACACAAATTTGTA 1241
1136 ACTTAGCATCAGAAATTTGATAGTTGATGCTATGACGTTCCACGCGGACCAATCATCT 1195
1242 TCATTAAATGTTGGGCGATGAGAGATCTCTCAGTTGTGGAATGATGTCACATGCTTTA 1301
1196 TGGTGAATGCAATGGGCCCATCCATAGATCCAAAGTTATGGGAAGAGCCAGAAAATTTA 1255
1302 AACCTGAGAGGTTTGTATGTGAAGAGAGAGAGAAAAGTTGGTAGCATTTGGCATGGGAA 1361
1256 AGCCGAGAGGTTTGAGAAAAGAGAGAGGATTAAGAGTTGATGCCATTTGGGATGGAC 1315
1362 GAAGGCTTGGCAGGAGAACCATGGCTATGCAAAAGTGTGAGCTTTACTTTGGGATTTG 1421
1316 GACGATCTTGGCCTGGTTTCAAGGCTAGCTCAACGCTAGTGAATTTGGCTCTTGGATCAT 1375
1422 TGATTTCAATGTTTTCAGCTGGAACAGTAAGTGAAGGAAAAGCTTGATATGACAGAGA--- 1478
1376 TGGTCCAAATGTTTTCAGTGGGAGAGGGTTGAAGAAAAGTATCTAGACATGAGAAAGTG 1435
1479 ACAATTTGATCACTTTGCAAGGTTAAATTCATTTGGAGGCAATGTGCAAGGCTCGCCAC 1538
1436 AAAAAGGAACCAATGCAAAAGCTACATCGTTGCAAGCTATGTGTAAGCTCGTCCCA 1495
1539 TTGCCACTAAAATTTGAATTTAAATTAATAGTATTTTAT 1580
1496 TTGTCCTAAGGTTTTAGATGCTTCTCTCCATCACTTTTCAT 1537

RESULT 12
AY050849

LOCUS
DEFINITION

Arabidopsis thaliana putative cytochrome P450 protein (At3g28740)

AY050849

1612 bp mRNA linear PLN 18-SEP-2002

complete cds.

AY050849

GI:15292830

FLI CDNA.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 1612)

Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,

Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,

Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,

Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C.,

Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,

Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K.,

Arabidopsis Full Length cDNA Clones

Unpublished

2 (bases 1 to 1612)

Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Banno,F., Dale,J.M.,

Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L.,

Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L.,

Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,

Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C.,

Koesema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M.,

Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P.,

Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and

Theologis,A.

Direct Submission

Submitted (08-AUG-2001) Plant Gene Expression Center, 800 Buchanan

Street, Albany, CA 94710, USA

RIKEN Genomic Sciences Center (GSC) members carried out the

collection and clustering of RAPL cDNAs (RAPL cDNA : 'Riken

Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,

Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,

Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the

sequencing and annotation of the RAPL cDNAs: Yamada,K., Liu,S.X.,

Pham,P.K., Banh,J., Banno,F., Dale,J.M., Goldsmith,A.D.,

Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,

Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H.,

Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koesema,E., Lam,B.,

Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,

Southwick,A., Tracy,S.E., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGECC) and Seki,M. (RIKEN GSC) contributed equally to

this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGECC)

contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis

genome submitted to GenBank.

Location/Qualifiers

1. .1612

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="3"

/clone="RAPL08-11-J17 (R1122)"

/note="This clone is in a modified pBluescript vector

(FLC-1) as a BamHI/XhoI insert.

ecotype: Columbia"

1. .1612

/gene="At3g28740"

1

/genes="At3g28740"

2. .1531

/gene="At3g28740"

/codon_start=1

/evidence=experimental

FEATURES

source

gene

5'UTR

CDS

JOURNAL Patent: WO 0216655-A 2278 28-FEB-2002;
 The Scripps Research Institute (US) ; Syngenta Participations AG
 (CH)

FEATURES

source	Location/Qualifiers
1. .1494	/organism="Arabidopsis thaliana"
	/mol_type="genomic DNA"
	/db_xref="taxon:3702"
BASE COUNT	399 a 341 c 355 g 399 t
ORIGIN	
Query Match	20.8%; Score 386.6; DB 6; Length 1494;
Best Local Similarity	57.0%; Pred. No. 1.1e-76;
Matches 839; Conservative	0; Mismatches 599; Indels 33; Gaps 6;
Qy	90 TTTTCTCGTGTCTATTCTCGCGCGTGAAGTTGTGTTTCCAAAGCAGAAAAATTGAGAAACA 149
Db	29 TTCTCTTGTGTTCTCTCTCTCATTTAATTTGGAAGAACTCAAGCGAAGCAATC 88
Qy	150 TACACACAGGTCCTCC---TCCTCTTCCATAAATAGGAAACCTTAACTCTCTCGAACAGC 206
Db	89 TACCTCGAGTCGGCATGTGTCATTACCGGTGATTGGTCATCTCGCGCTTCTCAAACCAAC 148
Qy	207 CAATCCACCGTTTCTTCCACGCGATGTGGAACAGTAGTCGGCAACG-----TGGTTTCCC 260
Db	149 CGATTCAATCGCACATTCCTCTCCCTCTCTCAATCCCTAAACAATGCTCCGATCTTCTCCC 208
Qy	261 TCTGTTTCGGTTCACGTCGTGCGCGTGTGCATCTCTCTCCTCCAACAGCATACCAAGAATGCT 320
Db	269 TCACCAAAACGACGTGTACTGGCGAAGACACCAAACTTCATCTCTCGCTAAACAGCTTG 328
Qy	381 TCTACAAACAACACACCGTAGGTCCTGCTCCACGGCGAGGACATGGCGCAACTCTCCGCC 440
Db	329 CGTACGATTACACAAACCATGATCGCAGCTTCTACGGTGAACCACTGGGGGTAACTCCGCC 388
Qy	441 GCATCACGCCCTGGACGTCCTCTCCACGAGCGGTCCACTCTCTTCTCCGGAATCCCGGA 500
Db	389 GCATCGGCTCCGTGAGAAATTTCTCAATACCGTCTCAATAGCTTTTGTCTATTCGTGA 448
Qy	501 GCGACGAGACGAAGCGTCTGATCGAGAGGTTGGTCTCGCGCAAGAACTCGAACGAGGAAG 560
Db	449 AAGACGAGATCCGACGACTTGTGTTTGTCTTCTCGAAC-----TTTTTCACAAG 499
Qy	561 AGTTTGCAGATGGAGATAGTTTCGATGTTTCAACGACTTAATTAACAAACATTAAGA 620
Db	500 AGTTTGTGAAGTGGATATGAATCAATGTTATCTGACTTAACTTCAACAACTTTTAA 559
Qy	621 GGATGATATCGGGGAAGAGGTTTTCACGAGAGGAGATGAGATGAAGAAGCTTGTAGGAAG 680
Db	560 GAATGGTCCCGGAAAAAGTTACTACGAGAGACGGTGTG-----AGATGATCCGAGG 613
Qy	681 CGAGGAGTTTCAGAGAGACTGTGACAGAAATGTTTGGAACTCATTGGGCTTGGCTAAACAAG 740
Db	614 CTAAACGTTCCGCGCAGCTTATAGCGGATGTGGTGTGCTGTGCTGTTGAAACGCTG 673
Qy	741 GAGATCACTTGCCTTTCCTCAGGTGGTTCGATTTTTCAGAAATGTGGAGAGCGCTTAAAGA 800
Db	674 TTGATTACTTACCGGTTTTCGGTTGG-----TTTCAGATTACGAGACACCGGTTAAGA 727
Qy	801 GTATCAGTAAGAGGTPACGATTCCTTGAATAAGATCCTTCAATGAGAACCGTGGCCAGCA 860
Db	728 AGTTGGCGGTTAGGCTCCACGAGTTCTTTCGAAGGATTTGGTTGATGAGAAACGAGAAGCTGA 787
Qy	861 ATGACCGCGAATTTCCATGATCGATCATCTCTCAAACTGCAAGAGACCCAGCCTCAGT 920
Db	788 AGGAGAAAGGAAACACTATGATCGATCACTTGTCTTACTCTGCAAGAAATCCCAACCGGATT 847
Qy	921 ACTACACTGACCAAAATCATAAAGCCCTTGCTCTCGCCCATGCTTTTGTGTGGAACTGACT 980

Db 29 TTCTCTTGGTGTCTCTCTCAATTAATGGAATACTCAAGCGAAGACCAATC 88
QY 150 TACACACAGTCTCTCC---TCTCTTCCCATATAGGAACCTTAACCTCTCTCGAACAGC 206
Db 89 TACCTCGAGTCGGCATGATGATACCGGTGATGATCTCTCGCGCTTCTCAACACAC 148
QY 207 CAATCCACCGTTCCTTCCCAAGCATGTGCAAAACAGTACGGCAACG-----TGGTTTCCC 260
Db 149 CGATTTCATGCAATTCCTCTCCCTCTCTCAATCCCTCAACCAATGCTCGGATCTTCTCCC 208
QY 261 TCTGGTTCGGTTCACCTCTGGCGGTGTCATCTCTCTCCATCCCAACAGATACCAAGATGCT 320
Db 209 TCGGATTCGGTAACCGACTGTTTTCGTAACCTCGTCACTCGATCGCGGAGATGTT 268
QY 321 TCACCAAAACAGACGTTGCTTGGCCAAACCGGTACTCTCTCTCTCGGAAATAATCATCT 380
Db 269 TCACCAAAACAGACGTCGTACTGCGCAACAGACCAAACTTCATCTCTCGTAAACAGTTG 328
QY 381 TCTAACAAACACACCGTAGGTCTCTGTCCTCCACGGCGAGACCTGGGCAACCTCGGC 440
Db 329 CGTACGATTAACAACACCATGATCGCAGCTTCTCTACGGTGACCACTGGCGTAACTCGGC 388
QY 441 GCATCACCGCCCTGGAGCTCTCTCCACGAGCGCTCCACTCTCTCGGAACTCCGA 500
Db 389 GCATCGCTCCGTCGAGATATCTCCAATCACCCTCTCAATAGCTTTTGTCTATTTCGTA 448
QY 501 GCGACGAGACGAACGCTCTGATCAGAGGTTGGTCTGGCCAAAGAACTCGAACGAGGAAG 560
Db 449 AAGACGAGATCCGACGACTGTGTTTCGTTCTCTCGAAC-----TTTTCACAG 499
QY 561 AGTTTCGGCGAGTGGAGATGATGTTGATGTTCAACGACTTAATTAACAACATATGA 620
Db 500 AGTTTGTGAAGTGGATGATGAATCAATGTTATCTGACTTAACTTAACAACATTTAA 559
QY 621 GGATGATATCGGGAAGAGTGTTCAGGAGAGAGTGAAGAGAGTGAAGAG 680
Db 560 GAATGTGCGCGGAACGTTTACTACGGAGCGGTGTTG-----AGGATGATCCGGAGG 613
QY 681 CGAGGAGTTCAGAGACGCTGTGACAGAAATGTTGGAACCTCATGGCTTGGCTTAAACAGG 740
Db 614 CTAAACGTCGCGGAGCTTATAGCGATGCTGTGCTGTGCTGTTGGAACGCTG 673
QY 741 GAGATCACTGCTTCTCTCAGTGTGTTGATTTTTCAGAAATGAGAGAGCGCTTAAAG 800
Db 674 TTGATTACTTACCGGTTTTCGGTTGG-----TTTCAGATTACGAGACACGCGTTAAGA 727
QY 801 GTATCAGTAAGAGTACGATTCCTCTTGAATAGATCTTCTATGAGAACCGTGCAGCA 860
Db 728 AGTTGGCGGTAGGTCGAGAGTTCTTGAAGGATTTGGTTGATGAGAAACGAGAAAGCTA 787
QY 861 ATGACCGCCAGAAATTCATGATGATCATCTCTCAAACTGCAAGAGACCCAGCTCAGT 920
Db 788 AGGAGAAGAAACACTATGATGATGATCACTGTTTACTCTGCAAGATCCCAACCGAT 847
QY 921 ACTACACTGACCAATCATCAAGCGCTGCTCTGGCCATGCTTTTGTGGAACTGATC 980
Db 848 ACTTACCGATCGTATCAATTAAGGAAACATGCTCGCTTTGATCTAGCAGGACCGACA 907
QY 981 CATCACTGGACTTTAGAGTGTCTATCTAATTTATGATCAACCCAGAGGTTTCA 1040
Db 908 CATCAGCGGTTCGTTAGAAATGGCAATGTCGAACCGTGTGAAACCATCCGATGATTA 967
QY 1041 AGAAGGCAAGAGATGAATGGACACTCAAGTGGGCAAGACCGCTTGTGTTAAATGATCAG 1100
Db 968 ACAAGGCGAGAGATGAATCGATAGAAAGATAGTTTGAAGAGCTTATGATGATCAG 1027
QY 1101 ACCTTCCAAACCTTCATATCTTGAAGAGATCACTCTGAGACACTTAGGTTGACCCCC 1160
Db 1028 ATATCTCAAACTGCTTATCTCTCAAAACATTTGTCTGAAAGGTTGCGCTTTATCTCG 1087
QY 1161 CGGCCCCCAATTTCAATACCTCATGTCTCTCAGAAATATTAATTAAGGATTCATA 1220

Db 1088 CGGCTCCCATGCTTCTCTCTCAGCTTGGCTCGGAAGATTTGTAAGTTGCAAGATACGATA 1147
QY 1221 TCCACGAGACACAAATTTGTGATCATTAATGTTGGGGCATGCAAGAGATCTCTCAGTTGT 1280
Db 1148 TGGCGGTGGCAACGATATAATTTGACCAATGTTGGGCTATACACAGAGATCTCTCAGCTAT 1207
QY 1281 GGAATGATGCCACATGCTTTTAAACCTGAGAGGTTTGTGTTGGAAGGAGGAGAAAGT 1340
Db 1208 GGGATGATCCAAATGAGCTTCAAGCCAGAGAGGTTTGAAGAAAGAGGAGAGCTCAGAAGC 1267
QY 1341 TGGTAGCATTTGGCATGGAAGAGGGCTTCCCAAGGAGAACCCATGCTATGCAAGATG 1400
Db 1268 TAATCCCGTTTGGGTTAGAAAGAGGGCGTGTCTGGTCTGGACTGGCTCATCGGCTTA 1327
QY 1401 TCAGCTTTTACTTTGGGATTTGATTTCAATGTTTGACTGGAAACGAGTAAGTAGAGAAA 1460
Db 1328 TAAACCTGACTTTGGATCATTTGATTTGATTTGGAATGGG---GAAGATTTGGAGAAG 1384
QY 1461 AGCTTGATATGACAGAGAAACAATTTGGATCACTTGTCAAGGTTAATTTCCATTTGGAGGCA 1520
Db 1385 AAGTGATATGATGAGGCAAGGTTTACAATGCTTAAAGCCAAAGCCTTTTGAAGCCA 1444
QY 1521 TGTGCAAGGCTCGCCACTTGCCTCACTAAAT 1551
Db 1445 TGTGCAAGCAGCTCCCTCTGTTGTTAAAT 1475

RESULT 15
AX651453 1494 bp DNA linear PAT 22-MAR-2003
LOCUS AX651453
DEFINITION Sequence 249 from Patent WO03000898.
ACCESSION AX651453
VERSION AX651453.1 GI:29154271
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,
Karagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.
Plant genes involved in defense against pathogens
Patent: WO 03000898-A 249 03-JAN-2003;
Syngenta Participations AG (CH)
Location/Qualifiers
1. 1494
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"

BASE COUNT 399 a 341 c 355 g 399 t
ORIGIN

Query Match 20.8%; Score 386.6; DB 6; Length 1494;
Best Local Similarity 57.0%; Fred. No. 1.1e-76;
Matches 839; Conservative 0; Mismatches 599; Indels 33; Gaps 6;

QY 90 TTTTCTCTGTTCTATTCTCTCGCGTGAAGTTGTTTCCAAAGCAGAAATTTGAGAAACA 149
Db 29 TTCCTCTGTTCT 88
QY 150 TACCAACAGGTCCTCC---TCTCTCTCCATAATAAGAAACCTTAACCTCTCTCGAACAGC 206
Db 89 TACCTCCAGTCCGCGATGTTCAATTCGCGTGAATTTGATGCTCTCTCTCTCTCTCTCTCT 148
QY 207 CAATCCACCGTTCCTTCCAAACGATGTCGAACAGTTCCTTAACCTCTCTCGAACAGC 206
Db 89 TACCTCCAGTCCGCGATGTTCAATTCGCGTGAATTTGATGCTCTCTCTCTCTCTCTCTCT 148
QY 207 CAATCCACCGTTCCTTCCAAACGATGTCGAACAGTTCCTTAACCTCTCTCGAACAGC 260
Db 149 CGATTTCATGCAATTCCT 208
QY 261 TCTGGTTCGGTTCAGTCTGCGCGTGTGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 320
Db 209 TCCGACTCGGTAAACGACTCTGTTTTCGTGTAATCTGTCATCTGATCGCGGAGGATGTT 268

Qy	321	TCACCAACACGAGCTTGGCTTGGCCAAACCGGCTACCTTCTCTCTCGGGAATAATCATCT	380
Db	269	TCACCAAAAAACGAGCTGCTGTCGGGAACAGACCAAACTTTCATCTCGCTTAAACACGTTG	328
Qy	381	TCTACAAACAACACACCGTAGGCTCTGCTCCACGCGGAGCACTGCGCAACCTCGCC	440
Db	329	CGTAGCATTTACAAACCATGATCGAGCTTCTTACGGTGACCACTGCGGTAACTCGCC	388
Qy	441	GCATCACCGCCCTGGACGCTCCTCTCCACGACGCGCTCCACTCCTTCTCGGAATCCGGA	500
Db	389	GCATCGGCTCGTGCAGATATTCTTCCAATCACCGTCTCAATAGCTTTTGTCTATTGTA	448
Qy	501	CGACGAGACGAAGCGTCTGATCAGAGAGTTGGTGTGGCCCAAGAACTCGAACGGAAG	560
Db	449	AAGACGAGATCCGACGACTTGTGTCTTCTCGTCTTCTCGGAAC-----TTTTCACAAG	499
Qy	561	AGTTTGGCGAGTGGAGATTAGTTTCGATGTTTCAACGACTTAACTTACAAACAATATGA	620
Db	500	AGTTTGTGAAGTGGATATGAATCAATGTTATCTGACTTAACTTCAACAACATTTTAA	559
Qy	621	GGATGATATCGGGGAAGAGGTTTTTACGGAGAGAGAGTGAGATGAAGAAAGTTGAGGAAG	680
Db	560	GAATGGTGGCGGAACGTTACTACGGAGACGCTGTTG-----AGATGATCCGGAGG	613
Qy	681	CGAGGGAGTTCAGAGAGACTGTGACAGAAATGTTGGAACTCATGGGCTTGGCTTAAAGA	740
Db	614	CTAAACGTGTCGGCAGCTTATAGCGATGTGGTGGCTTGTGCTGTGCTGGAACGCTG	673
Qy	741	GAGATCACTTGGCCCTTTCCTCAGTGGTTCGATTTTTCAGATGTGGAGAAAGGCTTAAAGA	800
Db	674	TTGATTACTTACCGGTTTTTCCGCTGG-----TTTCAGATTACGACACCGGTTAAGA	727
Qy	801	GTATCAGTAGAGGTACGATTCCATCTTGTAATAGATCCTTTCATGAGAAACGCTGCCAGCA	860
Db	728	AGTTGGCGGTAGGCTCGACGAGTTCTTTCGAGAGATTGGTTGATGAGAAAGAGAGCTA	787
Qy	861	ATGACCCCGAGAATTCATGATCGATCATCTCTCAAACCTCGAAGACACCGCTCAGT	920
Db	788	AGGAGAAAGGAAACACATATGATCGATCACTTGCTTACTCTGCAAGAAATCCCAACGGATT	847
Qy	921	ACTACACTGACCAAAATCATCAAAGGCCCTTGTCTGGCCATGCTTTTGGTGGAACTGACT	980
Db	848	ACTTCACGSAATCGTATCATTTAAAGGAAACATGCTCGCTTTGATCTAGCAGGGAACGACA	907
Qy	981	CATCAACTGGGACTTTAGAGTGGTCAATTATCTAATTTATTGAATCACCCAGAGGTGTGA	1040
Db	908	CATCAGCGGTTACGTTAGATGGCAATGTGCAAGCTGTGGAACCATCCGGATGTATGA	967
Qy	1041	AGAAAGCAAGAGATGAATTTGACACTCAAGTGGGAACAAGACCGCTTGTAAATGAGTCAG	1100
Db	968	ACAAAGCGAGAGATGAATCGATAGAAGATAGGTTTAGACAGGCTTTATGATGAATCAG	1027
Qy	1101	ACCTTCCAAAACCTTCCATATCTTAGGAAGATCATCCTTGGAGACACTTAGGTTGATCCCC	1160
Db	1028	ATATCTCAAAACCTCGCTTATCTCCAAAACATTTGTCTGAAACGTTTCGGCTTTATCTCTG	1087
Qy	1161	CGGCCCCAAATTTCTAAATACCTCATCTGTCTTCAGAAAGATATTACAAATTTGAAGGATTCAATA	1220
Db	1088	CGGCTCCCATGCTTCTTCTCCTCAGCTTGGCTCGGAAGATTGTAAAGTTGCAGGATACGATA	1147
Qy	1221	TCCACGAGACACAATTTGTGATCAATTAATGGTTGGGGCATGACAGAGATCCTCAGTTGT	1280
Db	1148	TGCCGCTGGCAGCATACTATTGACCAATGTGTGGCTATATACACAGAGATCCTCAGCTAT	1207
Qy	1281	GGAATGATGCCATATGCTTTTAACTCAGAGGTTTGTGTGGAAGGAGAGAGAGAAAAGT	1340
Db	1208	GGGATGATCCAATGAGCTTCAAGCCACAGAGGTTTGAAGAAAGAGAGAGAGCTCAGAAGC	1267
Qy	1341	TGGTAGCATTTGGCATCGGAAGAAGGCGTTGCCACGAGGAAACCCATGCCTATGCAAGTG	1400
Db	1268	TAATGCCGTTTGGGTTAGGAAGAAGGCGGTCTCTGGTTCTGGAATCGCTCATCGGCTTA	1327
Qy	1401	TCAGCTTTTACTTTTGGGATTTGTTGATTCAATGTTTTTGTGCTGAAACGAGTAGTGAGGAAA	1460

[illegible]

Search completed: October 21, 2003, 04:46:58
Job time : 6733 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2003, 02:38:44 ; Search time 3601 Seconds
(without alignments)
12547.076 Million cell updates/sec

Title: US-09-868-546A-1
Perfect score: 1859
Sequence: 1 gaaacactgacagacagca.....aataaatttttactttc 1859
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estun:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	596.4	32.1	598	13	BQ786681 saq71h06.
2	582.6	31.3	589	12	BI321303 saf50c06.
3	573	30.8	573	13	BUS79309 saf56f09.
4	491.8	26.5	600	13	BQ742415 saq43a09.

C	5	477.2	25.7	743	12	BI969319
	6	470.6	25.3	582	9	AW234443
	7	468	25.2	581	9	AW307234
	8	465.2	25.0	581	13	BU764794
	9	455.8	24.5	574	12	BI498223
	10	455.4	24.5	561	12	BM187538
	11	451.2	24.3	544	13	BQ079433
	12	444.6	23.9	564	13	BQ453555
	13	426	22.9	564	13	BU762419
	14	417	22.4	425	12	BI424885
	15	410.8	22.1	426	12	BI426262
	16	410	22.1	552	9	AW782116
	17	406.4	21.9	420	13	BQ742929
C	18	400.8	21.8	651	9	AW351157
	19	400.2	21.5	503	9	AI495626
	20	398.2	21.4	524	12	BI971995
	21	396.6	21.3	696	14	CB891448
	22	395.2	21.3	473	10	BG046583
C	23	383.6	20.6	666	9	AW309826
	24	374.6	20.2	789	9	AW102198
C	25	373.8	19.8	846	14	CB893434
	26	367.8	19.8	846	14	CB893434
	27	362.6	19.5	845	14	CB893246
C	28	355.6	19.1	608	9	AW309498
	29	346.2	18.6	610	13	BQ612582
	30	342	18.4	422	13	BQ080894
	31	337.4	18.1	491	13	BU762460
C	32	329	17.7	787	10	BG644813
	33	328.8	17.7	784	14	CB893907
	34	328.2	17.7	673	10	BG045616
C	35	321.8	17.3	754	12	BM779857
	36	318.6	17.1	770	10	BG648962
	37	316.4	17.0	769	14	CB893509
	38	314.4	16.9	576	13	BU764619
	39	313	16.8	593	10	BE205034
	40	308.8	16.6	596	9	AW185361
	41	305.4	16.4	587	10	BE190273
C	42	301.8	16.2	797	14	CA922487
	43	297	16.0	712	9	AW775904
	44	290.2	15.6	723	14	CB893510
C	45	288.2	15.5	591	13	BU545044

ALIGNMENTS

RESULT 1
BQ786681

LOCUS saq71h06.y1 Gm-cl076 Glycine max cDNA clone SOYBEAN CLONE ID:
DEFINITION Gm-cl076-5363 5', similar to TR:Q9ZWF2 Q9ZWF2 CYTOCHROME P450. 1,
mRNA sequence.

ACCESSION BQ786681
VERSION BQ786681.1 GI:21995153
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max

REFERENCE 1 (bases 1 to 598)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST Project

REFERENCE 1 (bases 1 to 598)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST Project

TITLE BQ786681 saq71h06.
JOURNAL BI321303 saf50c06.
COMMENT BUS79309 saf56f09.
BQ742415 saq43a09.

598 bp mRNA linear EST 26-JUL-2002
saq71h06.y1 Gm-cl076 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl076-5363 5', similar to TR:Q9ZWF2 Q9ZWF2 CYTOCHROME P450. 1,
mRNA sequence.
BQ786681
BQ786681.1 GI:21995153
EST.
Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccu@resgen.com web site: www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 421.

Location/Qualifiers

1. .598

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-c1076-5363"

/tissue_type="wounded cotyledons"

/dev_stage="11 day old seedlings"

/lab_host="DH10B"

/clone_lib="Gm-c1076"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from 11 day old seedlings treated with that were treated with 2 ugs/ml of a crude glucan elicitor preparation isolated from the mycelial walls of Phytophthora sojae. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. Plant material was provided by Michael G. Hahn (Complex Carbohydrate Research Center, University of Georgia) and the library was constructed by Anu Khanna (Lila Vodkin lab the University of Illinois)." 139 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 32.1%; Score 596.4; DB 13; Length 598;

Matches 597; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GAAACACTGACAGACGATAGTCTCTGGTGCAAGAAATCAATTGACCAAGCATGGGAAT 60

Db 1 GAAACACTGACAGACGATAGTCTCTGGTGCAAGAAATCAATTGACCAAGCATGGGAAT 60

Qy 61 GTTCTGGTGTGCTCTCATACGCTGTCCTTTTCTGTTCTATCTCTCGGCGTGAAGTT 120

Db 61 GTTCTGGTGTGCTCTCATACGCTGTCCTTTTCTGTTCTATCTCTCGGCGTGAAGTT 120

Qy 121 TGTCTTCCAAAGCAGAAATTGAGAAATACACACAGGTCCTCTCTCTCTTCCCATAT 180

Db 121 TGTCTTCCAAAGCAGAAATTGAGAAATACACACAGGTCCTCTCTCTTCCCATAT 180

Qy 181 AGGAAACCTTAACTCTCTGAAACAGCAATCCACCGTTCTTCCAAAGCATGCGAAACA 240

Db 181 AGGAAACCTTAACTCTCTGAAACAGCAATCCACCGTTCTTCCAAAGCATGCGAAACA 240

Qy 241 GTACGGCAACGTGTTTCCCTCTGTTGGTTCAGTCTGCGGTGCGGTGTCATCTCTCTCC 300

Db 241 GTACGGCAACGTGTTTCCCTCTGTTGGTTCAGTCTGCGGTGCGGTGTCATCTCTCTCC 300

Qy 301 AACAGCATACCAAGAAGTCTTCCAAACAGCAGGTTGCCCTTGGCCAAACGGCTACCTTC 360

Db 301 AACAGCATACCAAGAAGTCTTCCAAACAGCAGGTTGCCCTTGGCCAAACGGCTACCTTC 360

Qy 361 TCTCTCGGAAATACATCTTCTACAAACACACACCGTAGGCTCTGCTCCACCGCGCA 420

Db 361 TCTCTCGGAAATACATCTTCTACAAACACACACCGTAGGCTCTGCTCCACCGCGCA 420

Qy 421 GCATGGCGCAACCTCCGCGCATACCGCCCTCTGACGCTCTCTCCACGAGCGGTCCA 480

Db 421 GCATGGCGCAACCTCCGCGCATACCGCCCTCTGACGCTCTCTCCACGAGCGGTCCA 480

Qy 481 CTCCTTCTCCGAATCCGAGCGAGACGAGCAAGAGCTCTGATGCGAGAGTTGGTGTGGC 540

Db 481 CTCCTTCTCCGAATCCGAGCGAGACGAGCAAGAGCTCTGATGCGAGAGTTGGTGTGGC 540

Qy 541 CAAGAACTCGAAGAGAGAGTTTGGCGAGTGGAGATTAGTTCGATTTCAACGAC 598

Db 541 CAAGAACTCGAAGAGAGAGTTTGGCGAGTGGAGATTAGTTCGATTTCAACGAC 598

RESULT 2

BI321303

LOCUS

DEFINITION

sa50c06.v3 Gm-c1077 Glycine max cDNA clone GENOME SYSTEMS CLONE

ID: Gm-c1077-2027 5' similar to TR:Q9ZWF2 Q9ZWF2 CYTOCHROME P450.

1. mRNA sequence.

ACCESSION

BI321303

VERSION

BI321303.1

KEYWORDS

EST.

SOURCE

Glycine max (soybean)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

REFERENCE

1 (bases 1 to 599)

AUTHORS

Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V., Khanna

, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,

Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers

, Y., Person, B., Swallier, T., Gibbons, M., Pape, D., Harvey, N., Schurk

, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann

, R., Waterston, R. and Wilson, R.

Public Soybean EST Project

Unpublished

CONTACT: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact via email: ccu@resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 421.

Location/Qualifiers

1. .589

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1077-2027"

/tissue_type="18 day old 'Williams' seedlings"

/dev_stage="18 day old 'Williams' seedlings"

/lab_host="DH10B"

/clone_lib="Gm-c1077"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The mRNA was isolated from cotyledons of 18-day-old

'Williams' seedlings which were greenhouse grown in

potting soil. The cotyledons were flash-frozen in liquid

nitrogen. Stratagene's cDNA Synthesis Kit (catalog number

200401) was used to synthesize the cDNA. First-strand

synthesis was performed with 5-methyl dCTP, hence the

ligated cDNA was hemimethylated. A modification of

Stratagene's first-strand synthesis primer was used. An

'anchor' nucleotide (V=A, C, or G) was added to the 3' end

of the primer (GAGAGAGAGAGAGAGAGTCTCGAG(T)18V) to

anchor the primer at the 5' end of the poly(A) tract.

After second-strand synthesis, the cDNA ends were filled

in with cloned Pfu DNA, ligated to EcoRI adapters and

subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The xhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (40U/ul); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using Sephacryl 8-500 High Resolution (Pharmacia Biotech) in a 2-mm diameter column and a bed volume of approximately 1ml. The column eluent was precipitated, redissolved, and ligated into Stratagene's pBluescript II XR predigested vector (pBluescript II SK(+)) vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). This library was constructed in the laboratory of Dr. Paul Keim and Dr. Virginia H. Coryell at Northern Arizona University."

BASE COUNT 173 a 127 c 134 g 155 t
ORIGIN

Query Match 31.3%; Score 582.6; DB 12; Length 589;
Best Local Similarity 99.3%; Pred. No. 2.6e-120;
Matches 585; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 748 CTTCGCTTCCCTCAGGTGGTTCGATTTTCAGATGTGGAGAGCGCTTAAAGAGTATCAG 807
DB 1 CTTCGCTTCCCTCAGGTGGTTCGATTTTCAGATGTGGAGAGCGCTTAAAGAGTATCAG 60

QY 808 TAAGAGGTACGATTCATCTTGAATAAGATCTCTTCATGAGAACCGTGCCAGCAATGACCG 867
DB 61 TAAGAGGTACGATTCATCTTGAATAAGATCTCTTCATGAGAACCGTGCCAGCAATGACCG 120

QY 868 CAGAAATTCATGATCGATCATCTCTCAAATGCAAGAGAGCCAGCCCTCAGTACTACAC 927
DB 121 CCAGAAATTCATGATCGATCATCTCTCAAATGCAAGAGAGCCAGCCCTCAGTACTACAC 180

QY 928 TGACCAAAATCATCAAAAGGCTTGCTCTGCGCATGCTTTTGTGGAACTGACTCATCAAC 987
DB 181 TGACCAAAATCATCAAAAGGCTTGCTCTGCGCATGCTTTTGTGGAACTGACTCATCAAC 240

QY 988 TGGGACTTTAGAGTGTCTATTATCTAATTTATGTAATCACCAGAGGTGTGAAGAAGGC 1047
DB 241 TGGGACTTTAGAGTGTCTATTATCTAATTTATGTAATCACCAGAGGTGTGAAGAAGGC 300

QY 1048 AAGAGATGAATGGACATCAAGTGGGAGCAAGACCGCTGTTAAATGATCAGACTTCC 1107
DB 301 AAGAGATGAATGGACATCAAGTGGGAGCAAGACCGCTGTTAAATGATCAGACTTCC 360

QY 1108 AAAAATTTCCATATCTTTAGGAAGATCATCTTCGAGACATTAGGTGTACCCCGGCCCC 1167
DB 361 AAAAATTTCCATATCTTTAGGAAGATCATCTTCGAGACATTAGGTGTACCCCGGCCCC 420

QY 1168 AATTCTTAATACCTCATGTGCTCTTCAAGATATTCAATTTGAAGGATTCATATCCACG 1227
DB 421 AATTCTTAATACCTCATGTGCTCTTCAAGATATTCAATTTGAAGGATTCATATCCACG 480

QY 1228 AGACAAATTTGATCATTAATGTTGGGCGATGACAGAGATCTCTCAGTTGGAATGA 1287
DB 481 AGACAAATTTGATCATTAATGTTGGGCGATGACAGAGATCTCTCAGTTGGAATGA 540

QY 1288 TGCCACATGCTTTAAACCTGAGAGGTTTGTATGTTGGAGAGGAGGAGAA 1336
DB 541 TGCCACATGCTTTAAACCTGAGAGGTTTGTATGTTGGAGAGGAGGAGAA 589

RESULT 3
BUS79309
LOCUS
DEFINITION
sars6f09.y1 Gm-cl074 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-cl074-6425 5' similar to TR:Q9ZRW6 Q9ZRW6 CYTOCHROME P450 ;,
mRNA sequence.
ACCESSION
BUS79309
VERSION
BUS79309.1 GI:23064536
KEYWORDS
EST.

SOURCE ORGANISM

Glycine max (soybean)
Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosoids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE

1 (bases 1 to 573)

AUTHORS

Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.

TITLE

Public Soybean EST Project

JOURNAL

Unpublished

COMMENT

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: cu@resgen.com web site:
www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 425.

Location/Qualifiers

FEATURES source

1..573
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl074-6425"
/tissue_type="seedlings induced for HR (hypersensitive response)"
/dev_stage="9-11 day old"
/lab_host="DH10B"
/clone_lib="Gm-cl074"

XhoI; The cDNA library was constructed from mRNA isolated from 9-11 day old seedlings that were induced for HR (hypersensitive response) by vacuum infiltrating plant tissue with *Pseudomonas syringae* pv. *glycinia* carrying the avrB gene (Genetics 141:1597-1604). Plant tissue (expanded unifoliate leaves) was collected at 2, 4, 8, 12, 24, 36, and 53 hrs after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. Plant care, inoculations, and library construction were performed by Steve Clough (Lila Vodkin lab, University of Illinois)."

BASE COUNT 173 a 99 c 135 g 166 t

ORIGIN

Query Match 30.8%; Score 573; DB 13; Length 573;
Best Local Similarity 100.0%; Pred. No. 3.7e-118;

Matches 573; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1054 TGAATTGGACACTCAAGTGGGAGCAAGCCGCTTGTAAATGAGTCAGACCTTCCAAAAT 1113
DB 1 TGAATTGGACACTCAAGTGGGAGCAAGCCGCTTGTAAATGAGTCAGACCTTCCAAAAT 60

QY 1114 TCCATATCTTAGGAAGATCATCTCTTGAGACACTTAGGTGTGACCCCGGCCCAATTCT 1173

Db 541 CGAGTAAGTGAAGAAGCTTGATATGACAGAGAAATATGGATCACCTTGTCAAGTTA 600

RESULT 5

BI969319/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI969319 743 bp mRNA linear EST 23-OCT-2001
 GM830008A10A12 Gm-r1083 Glycine max cDNA clone Gm-r1083-2711 3',
 mRNA sequence.

BI969319 1 GI:16343724
 EST.
 Glycine max (soybean)

Glycine max
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

1 (bases 1 to 743)
 Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Coryell, V.,
 Erpelting, J., Raph, C., Shoop, E., Fardinas, J., Liu, L. and Lewin, H.
 A Functional Genomics Program for Soybean (NSF 9872565)

Other ESTs: AW707100 corresponding to Gm-cl028-2839 (5')
 Contact: Vodkin, L.O., PI, A Functional Genomics Program for
 Soybean (NSF 9872565)

Lewin, H. A., Director, Keck Center for Comparative and Functional
 Genomics
 University of Illinois

Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA
 Tel: (217) 244-6147
 Fax: (217) 333-4582

Email: l-vodkin@uiuc.edu
 This clone is available through: Incyte Genomics, 4633 World
 Parkway Circle St. Louis, Missouri 63134. Phone (800) 430-0030 or
 (314) 427-3222 FAX: (314) 427-3324. Web site:
 http://www.incyte.com/reagents/catalog.jsp?page=clones/collaboratio
 n/index

Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'.
 Location/Qualifiers

1. .743
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="Gm-r1083-2711"

/clone_lib="Gm-r1083"
 /notes="The library Gm-r1083 is a sequence-driven, reracked
 set of 4,992 clones selected from cDNA libraries from
 various tissues and stages of development of soybean. It
 represents 1117 sequences from the progenitor library
 Gm-cl009 (from mature roots of 2 month old greenhouse
 grown 'Williams' soybean plants); 820 sequences from the
 progenitor library Gm-cl013 (from 2 to 3 week old whole
 plants of Williams); and 3055 sequences from library
 Gm-cl028 (from 'Supernod' plants whose seedlings were
 inoculated with Bradyrhizobium japonicum, courtesy of Dr.
 Gary Stacey). The 5' ESTs of the source clones from the
 different progenitor libraries was used to select
 singletons, or a representative of each contig, which were
 reracked to form library Gm-r1083. The cDNA clones of the
 reracked Gm-r1083 library were then sequenced at the 3'
 end. The contig analysis to select unique genes was
 performed by the laboratory of Ernest Retzel, Center for
 Computational Genomics and Bioinformatics, University of
 Minnesota, http://web.ahc.umn.edu/biodata/nsfsoy/.

Reracking was performed by Incyte Genomics, St. Louis,
 http://www.incyte.com, and 3' sequencing by the Keck
 Center for Comparative and Functional Genomics, University
 of Illinois, http://www.lie.uiuc.edu/biotech/keck.html.

Note: The corresponding 5' EST from each clone in the
 Gm-r1083 library is listed in the 'OTHER EST' field. The
 detailed information on the source library for each clone
 can also be obtained by referring to the Incyte Genomics
 clone ID of the original cDNA library that is also listed

FEATURES

source

BASE COUNT	201 a	171 c	116 g	226 t	29 others
ORIGIN	under 'OTHER EST'."				
Query Match	25.7%; Score 477.2; DB 12; Length 743;				
Best Local Similarity	83.1%; Pred. No. 1.5e-96;				
Matches 571; Conservative	0; Mismatches 96; Indels 20; Gaps 3;				
QY	988	TGGGACTTTAGAGTGGTCAATTCATTAATTTATTTGAATCACCAGAGGTTTGAAGAAGC	1047		
DB	743	TNNNNNTTTAGAGTGGGCATTANCAANNAGTGAATCANCNNGAGGTGCTCAGAANN	684		
QY	1048	AAGAGATGAATTTGGACACTCAAGTGGGACAAGACCGCTTTTAATAGTACAGACTTCC	1107		
DB	683	ANGAGATGAGTTGGACGCTCAAGTAGGACCATCANNNTGTTAAATGAGTACAGACTTCC	624		
QY	1108	AAACCTTCCATATCTTAGGAAGATCATCTCTTGAGACACTTAGTGTGTACCCCGGCCCC	1167		
DB	623	NNNACTTCTTATCTCAGAAAGATAGTCTTGAACACATTAGTGTGTACCCCTCCGCTCC	564		
QY	1168	AATCTTAATACCTCATGTGTCTTCAGAAAGATATTAACAATTGAAGGATTCAATATCCCAG	1227		
DB	563	AATCTTAATACCAACAGCTGGCTTCAGAAAGATCAATATCGAAGGATTCAATGTCCAGC	504		
QY	1228	AGACACAATTGTGATCAATTAATGTTGGGGATGACAGAGATCTCAGTTGTGGAATGA	1287		
DB	503	AGACACAATTGTGATCAATTAATGTTGGGGATGACAGAGATCTCAGTTGTGGAATGA	444		
QY	1288	TGCCACATGCTTTAAACCTGAGAGGTTGATGTGGAAGGAGAGAGAAAGCTTGGTAGC	1347		
DB	443	TGGCACAAGCTTTAAACCTGAGAGGTTTGTGAAGAGAGAGAGAGAGAGTGGTAGC	384		
QY	1348	ATTGGGCATGGGAAGAAGGGCTTGGCCAGGAGAACCCATGGCTATGCAAAAGTGCAGCTT	1407		
DB	383	ATTGGTATGGGAAGAAGGGCTTGGCCAGGAGAACCCATGGCTATGCAAAAGTGTAGCTA	324		
QY	1408	TACTTTGGGATTTGATTCAATGTTTGTGCTGGAACGAGTAACTAGTGAAGAAAGCTTGA	1467		
DB	323	TACTTTGGGATTTGATTCAATGTTTGTGCTGGAACGAGTAACTAGTGAAGAAAGCTTGA	264		
QY	1468	TATGACAGAGAACCAATTGGATCACCTTCTCAAGGTTTAATTCATTTGGAGGCCATGTCAA	1527		
DB	263	TATGACAGAGATTAATTTGGATCACCTTCTCAAGGTTTAATTCATTTGGAGGCCATGTCAA	204		
QY	1528	GGCTCGCCCACTTGGCCATTAATAATTTGAATTAATTAATTAATTAATTTATTTGGTAA	1587		
DB	203	AGCCCGCCCACTTGGCCCAAAAGTTGAAGTTATTAACAA-----TATTTATTTGGTAT	149		
QY	1588	ACTTTGGGTGATTCAGAACTTAATTAATTTAGTGTGTGAAGTGGTGTATATATA	1647		
DB	148	ATTTGGGTG-----AGGATCTAATTAATTTGGGTGTGTAAG-----TCTA	104		
QY	1648	TACATTTCAAATTAATTAATTTGTC	1674		
DB	103	TGCATTTAAATTAATTAATTTGTC	77		
RESULT 6	582 bp mRNA linear EST 02-DEC-2001				
AW234443	Gm-cl028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:				
LOCUS	sf25c03.yl Gm-cl028-989 5' similar to TR:Q9ZWF2 Q9ZWF2 CYTOCHROME P450. ;				
DEFINITION	mRNA sequence.				
ACCESSION	AW234443				
VERSION	AW234443.1 GI:6566811				
KEYWORDS	EST.				
SOURCE	Glycine max (soybean)				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.				
REFERENCE	1 (bases 1 to 582)				


```

/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl009-3792"
/lab_host="Xl10-Gold"
/clone_lib="Gm-cl009"
/notes="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:
XhoI. The mRNA was isolated from entire roots of
2-month-old 'Williams' plants that were greenhouse grown
in 5-gallon pots. To suppress nodulation, Black Gold
All-Purpose potting soil was supplemented with: 0.36g/L
available phosphoric acid (P2O5), 20mg/L urea N, 0.16g/L
S, 0.49mg/L B, 2.5mg/L Cu, 0.15g/L Fe, 13.53mg/L Mn,
0.26mg/L Mo, 14mg/L Zn, 20mg/L Ca, and the following
nutrients in a slow-release form (Osmocote): 0.165g/L
ammonia N, 0.185g/L nitrate N, 0.35g/L available
phosphoric acid, and 0.35g/L soluble potash. No nodules
were visible on the roots at harvest. Stratagene's cDNA
Synthesis Kit (catalog #200401) was used to synthesize the
cDNA. First-strand synthesis was performed with 5-methyl
dCTP, hence the ligated cDNA is hemimethylated.
Stratagene's first-strand synthesis primer was used
[GAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)-18]. After
second-strand synthesis is, the cDNA ends were 'polished'
with clone Pfu DNA polymerase, ligated to EcoRI adapters,
and phosphorylated. The XhoI site within the first-strand
synthesis primer was restricted by digestion with XhoI;
all XhoI sites in the cDNA would be protected by their
hemimethylated status. The cDNA constructs were
size-fractionated with a 400bp cutoff, using a SizeSep 400
Spin column from Pharmacia. The column eluent was then
ligated into Stratagene's pBluescript II XR predigested
vector (pBluescript II SK(+)) that had been digested with
EcoRI and XhoI, and phosphorylated. Both the white and
blue colonies appear to contain recombinant plasmids with
cDNA inserts. This library was constructed by Dr. Paul
Keim and Dr. Virginia Corvelli." 1 others
BASE COUNT 185 a 103 c 149 g 143 t
ORIGIN
Query Match 25.2%; Score 468; DB 9; Length 581;
Best Local Similarity 87.8%; Pred. No. 1.6e-94;
Matches 510; Conservative. 0; Mismatches 71; Indels 0; Gaps 0;
QY 903 AAGAGACCCAGCTCAGTACTACACTGACCAATCATCAAGGCCCTTGTCTGGCCATGC 962
Db 1 AAGAGACACAGCGCTGACTATTATACCGCAAAATCATCAAGGACTTGTCTGGCTATGC 60
QY 963 TTTTGGTGAAGTCACTCATCACTGGGACTTTAGAGTGGTCAATTCATTAATTTATGA 1022
Db 61 TCTTTGGCGGACAGACTCGTCACTGGAACTTTAGAGTGGCAATTAATTTATGA 120
QY 1023 ATCACCAGAGGTGTTGAAGAGGCAAGAGATGAATTGGACACTCAAGTGGGCAAGACC 1082
Db 121 ATGACCCAGAGGTGCTGCAAGAGGCAAGAGATGAGTTGGACCTCAAGTAGGACCATC 180
QY 1083 GCTTGTAAATGATGATGACACCTTCCAAATCTTCATATCTTAGGAAGATCATCTTGAGA 1142
Db 181 GCGTGTAAATGATGATGACACCTTCCAAATCTTCATATCTAGGAAGATGATCTTGAAA 240
QY 1143 CACTTAGGTGTACCCCGCCGCCCAATTCATACTCTATCTGCTCAGAGAATATTA 1202
Db 241 CACTTAGGTGTACCCCTCGGCTCCAAATTCATATACCACAGCTGGCTTCAGAAGACATCA 300
QY 1203 CAATTGAAGGATTCATATCCACGAGACACAAATTTGTGATCATTAATGGTTGGGGCATGC 1262
Db 301 ATATCGAAGGATTCATGTTCCAGGACGACAAATTTGTGATTAATGGTTGGGGCATGC 360
QY 1263 AGAGAGATCCTCAGTTGTGGAATGATGCCACATGCTTTAAACCTGAGAGGTTTGTGTTGG 1322
Db 361 AAAGAGATCCTAAGATATATGAAGATGCGCAAGCTTTTAAACCTGAGAGGTTTGTGTTGA 420
QY 1323 AAGGAGAGGAGAAAAGTTGGTAGCATTTGGCATGGGAGAGAGGCTTCCCGCAGAGAAC 1382

```

```

Db 421 AAGGAGAGGAGAGAGTGGTAGCATTTGGTATGGGAAGATGGCTTGGCCAGAGAAC 480
QY 1383 CCATGGCTTATCAAAAGTGTCTAGCTTTACTTTGGGATTTGATTCATCAATGTTTGTGCTGA 1442
Db 481 CCATGGCTTATGCAAGTGTAGCTATATACTGTGGGATTAATGATTCATCAATGTTTGTGCTGA 540
QY 1443 AACGAGTAAAGTGAAGAAAGCTTATATATGACAGAGAACAAAT 1483
Db 541 AACGAGTGAAGTGAAGAAAGCTTATATATGACAGAGAACAAAT 581

RESULT 8
BU764794 581 bp mRNA linear EST 10-OCT-2002
LOCUS sa006f04.y2 Gm-cl080 Glycine max cDNA clone SOYBEAN CLONE ID:
DEFINITION Gm-cl080-2960 5' similar to TR:Q9XF0 Q9XF0 CYTOCHROME P450
MONOOXYGENASE ;, mRNA sequence.
ACCESSION BU764794
VERSION BU764794.1 GI:23733217
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 581)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,B., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:
www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 420.
Location/Qualifiers
1..581
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl080-2960"
/tissue_type="Roots of 8 day old 'Bragg' supernodulating
mutant NTS382 seedlings"
/dev_stage="8 days old"
/lab_host="DH10B"
/clone_lib="Gm-cl080"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The mRNA was isolated from roots of 8 day old
'Bragg' supernodulating mutant NTS382 seedlings that were
infected with Bradyrhizobium japonicum, strain USDA 110, 72
hours prior to harvest. Dr. Gary Stacey generously donated
the tissue. The roots were flash-frozen in liquid
nitrogen. Stratagene's cDNA Synthesis Kit (catalog number
200401) was used to synthesize the cDNA. First-strand
synthesis was performed with 5-methyl dCTP, hence the
ligated cDNA was hemimethylated. A modification of
Stratagene's first-strand synthesis primer was used. An
'anchor' nucleotide (V=A, C, or G) was added to the 3' end
of the primer [GAGAGAGAGAGAGAGAACTAGTCTCGAG(T)18] to
anchor the primer at the 5' end of the poly(A) tract.
FEATURES
source

```


Query Match 24.5%; Score 455.8; DB 12; Length 574;
Best Local Similarity 86.9%; Pred. No. 8.9e-92;
Matches 499; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 826 CTTGATAGATCCCTCATGAGAACCGTCCGACGATGACCGCCAGAAATCCATCATCGA 885
DB 1 CTTGATAGATCCCTCATGAGAACCGTAAACACAGGACCGGAGAAATCCATCATCGG 60

QY 886 TCATCTCTCCCAAACTGCAAGAGACCCAGCGCTCAGTACTACACTGACCAATCATCAAAGG 945
DB 61 TCATCTCTCCCAAACTGCAAGAGACACAGCGCTGACTATATACCGACCAATCATCAAAGG 120

QY 946 CTTGCTCTGGCATGCTTTTGGTGGAACTGACTCATCAACTGGGACTTTAGATGGTC 1005
DB 121 ACTTGTCTTGGCTATGCTCTTGGCGGAAACAGACTGCTCAACTGGAACTTTAGATGGGC 180

QY 1006 ATTATCTAATTTATTGAATCAACCCAGAGGTGTGAAGAGGCAAGAGATGAATGGACAC 1065
DB 181 ATTATCTAATTTAGTGAATGACCCAGAGGTGTGCAAGAGGCAAGAGATGATGGACGC 240

QY 1066 TCAAGTGGGACAGACCGCTGTGTAATGAGTCAGACCTTCCAAACTTCCATATCTTAG 1125
DB 241 TCAAGTAGGACAGATCGCTGTGTAATGAGTCAGACCTTCCAAACTTCCATATCTCAN 300

QY 1126 GAAGATCATCTTGGACACTTAGGTTGTACCCCGCCGCCCAATTTCTAATACCTCATGT 1185
DB 301 GAAGATAGTCTTGTGAACACTTAGGTTGTACCCCGCTCCCAATTTCTAATACCACTGT 360

QY 1186 GTCTTCAGAGATATTACAAATGAAGGATTCATATCCACGAGACACAAATTTGTGATCAT 1245
DB 361 GGCTTCAGAGACATCAATATCGAAGATTCATATTCACGAGACACAAATTTGTGATCAT 420

QY 1246 TAATGTTGGGCGATCGAGAGATCCTCAGTTGTGGAATGATGCCACATGCTTTAAACC 1305
DB 421 TAATGTTGGGCGATCGAAGAGATCCTAAGATATGGAAGATGCGACAGCTTTAAACC 480

QY 1306 TGAGAGTTTGTGTCGAGGAGAGAGAGAAAGTTGGTAGCATTTGGCATGGAAGAG 1365
DB 481 TGAGAGTTTGTGTCGAGGAGAGAGAGAGAGAGTTGGTAGCATTTGGTAGGAGAGAG 540

QY 1366 GGCTTCCCGAGAGAACCCATGGCTATGCAAGT 1399
DB 541 GGCTTCCCGAGAGAAACCCATGGCTATGCAAGT 574

RESULT 10
BM187538
LOCUS
DEFINITION
saj86a06.y1 Gm-c1074 Glycine max cDNA clone SOYBEAN CLONE ID:
Gm-c1074-3467 5' similar to TR:Q9XFX0 Q9XFX0 CYTOCHROME P450
MONOOXYGENASE ; mRNA sequence.

ACCESSION
BM187538
VERSION
BM187538.1 GI:17518496
KEYWORDS
Glycine max (soybean)
SOURCE
Glycine max
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.

REFERENCE
1 (bases 1 to 561)
Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers
Y., Person, B., Swaller, T., Gibbons, M., Pade, D., Harvey, N., Schurk
R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Willson, R.
Public Soybean EST Project
Unpublished

TITLE
JOURNAL
COMMENT
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:
www.resgen.com

Seq primer: -40RP from Gibco
High quality sequence stop: 425.

FEATURES

Location/Qualifiers
1..561
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-c1074-3467"
/tissue_type="seedlings induced for HR (hypersensitive response)"
/dev_stage="9-11 day old"
/lab_host="DH10B"
/clone_lib="Gm-c1074"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from 9-11 day old seedlings that were induced for HR (hypersensitive response) by vacuum infiltrating plant tissue with Pseudomonas syringae pv. glycinea carrying the avr gene (Genetics 141:1597-1604). Plant tissue (expanded unifoliate leaves) was collected at 2, 4, 8, 12, 24, 36, and 53 hrs after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. Plant care, inoculations, and library construction were performed by Steve Clough (Lila Vodkin lab, University of Illinois)."

BASE COUNT 172 a 117 c 130 g 142 t

ORIGIN

Query Match 24.5%; Score 455.4; DB 12; Length 561;
Best Local Similarity 88.2%; Pred. No. 1.1e-91;
Matches 495; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 711 TGTGGAACTCATGGCTTGGCTAAAGGAGATCACTTCCTTCCTCAGGTGGTCG 770
DB 1 TGTGCAACTCATGGGTTGGCTAACCAAGGAGATCACTTACCTTCTCAGGTGGTCG 60

QY 771 ATTTTCAGATGTGGAGAGCGCTTAAGAGTATCAGTAAGAGGTACCATCTTGA 830
DB 61 ATTTTCAGAGCTGGAGAGAGGTTGAAGATATCAGTAAGAGGTATGATACCATCTTGA 120

QY 831 ATAAGATCCTTCATGAGAACCGTCCAGCAATGACCGCAGAAATTCATGATCGATC 890
DB 121 ATAAGATCCTTCATGAGAACCGTAAACACAGGACCGGAGAAATTCATGATGTCATC 180

QY 891 TCCTCAAACTGCAAGAGACCCAGCGCTCAGTACTACACTGACCAATCATCAAAGCGCTTG 950
DB 181 TCCTCAAACTGCAAGAGACACAGCGCTGACTATTATACCGACCAATCATCAAAGGACTTG 240

QY 951 CTCTGGCCATGCTTTTGGTGGAACTGACTCATCACTGGGACTTTAGAGTGGTCAATTAT 1010
DB 241 CTCTGGCTATGCTCTTTGGCGGAAACAGACTCGTCAACTGGAACCTTAGAGTGGGCAATTAT 300

QY 1011 CTAATTTATTGAATCACCACAGAGGTGTTGAAGAGGCAAGAGATGAATTTGACACTCAAG 1070
DB 301 CTAATTTAGTAATGACCACAGAGGTGCTGCAAGAGGCAAGAGATGAGTTGACCGCTCAAG 360

QY 1071 TGGGACAGACCGCTGTTAAATGAGTCAGACCTTCCAAAACCTTCATATCTTAGGAAGA 1130
 |||||
 Db 361 TAGGACAGATCGGCTGTTAAATGAGTCAGACCTTCCAAAACCTTCATATCTTAGGAAGA 420
 |||||
 QY 1131 TCATCTTGGAGACATTTAGTTGTACCCCGGGCCCAATCTTAATACCTCATGTGTCTT 1190
 |||||
 Db 421 TAGTCTTGAACAACATTTAGTTGTACCCCGGGCTCCCAATCTTAATACCAACAGTGGCTT 480
 |||||
 QY 1191 CAGAAGATATTACAAATTGAAGGATTCAATATCCACGAGACACAAATTTGATCATTAATG 1250
 |||||
 Db 481 CAGAGACATCAATATCGAAGATTCAATGTTCCACGAGACACAAATTTGATTAATG 540
 |||||
 QY 1251 GTTGGGCATCGAGAGATC 1271
 |||||
 Db 541 GTTGGGCATCGAAGAGATC 561
 |||||

RESULT 11

BQ079433 544 bp mRNA linear EST 04-APR-2002
 LOCUS san13f02.y1 Gm-cl084 Glycine max cDNA clone SOYBEAN CLONE ID:
 DEFINITION Gm-cl084-3724 5', similar to SW:CB1E_GLYEC P93147 CYTOCHROME P450
 81EL ;, mRNA sequence.

ACCESSION

BQ079433
 VERSION BQ079433.1 GI:19934403

KEYWORDS

EST.

SOURCE

Glycine max (soybean)

ORGANISM

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 544)

Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna

A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

,X., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk

,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann

,R., Waterston,R. and Wilson,R.

Public Soybean EST Project

Unpublished

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact: ccu@resgen.com web site:

www.resgen.com

Trace considered overall poor quality

Seq primer: -40RP from Gibco

High quality sequence stop: 1.

Location/Qualifiers

1. 544

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl084-3724"

/tissue_type="Etiolated hypocotyls (Williams 82)"

/lab_host="DH10B"

/clone_lib="Gm-cl084"

/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:

XhoI; The cDNA library was constructed by M. Bhattacharyya

from mRNA isolated from etiolated hypocotyls from the

cultivar Williams 82. Tissue was inoculated with

Phytophthora soyae race 1 and tissues were harvested 2 and

4 hours following infection. The library is the pool of

these two time points. Complementary DNA was synthesized

from mRNA using a primer consisting of a poly(dT) sequence

with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by M. Bhattacharyya in the laboratory of Dr. Randy Shoemaker at Iowa State University."

BASE COUNT 179 a 109 c 111 g 145 t

ORIGIN

Query Match 24.3%; Score 451.2; DB 13; Length 544;
 Best Local Similarity 89.3%; Pred. No. 9.5e-91;
 Matches 486; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
 QY 703 GACAGAAATGTTGGAACTCATGGCTTGGCTTAAAGGAGATCACTTCCCTTCTCCTCAG 762
 |||||
 Db 1 GACACACATGCTAGAACTCATGGGGTGGCTTAAGGAGATCACTTACCTTCTCTAAG 60
 |||||
 QY 763 GTGGTTCGATTTTCAGAAATGTGGAGAGCGCTTAAAGAGATATCAGTAAGAGGTACGATTC 822
 |||||
 Db 61 GAGGATCGATTTTTCAGAAATGTGGAGAGCGCTTGAAGAGATATTAATAAAGGTATGATAC 120
 |||||
 QY 823 CATCTTGAATAGATCCTTTCATGACACCGTCCAGCAATGACCGCAGCAATTCATGAT 882
 |||||
 Db 121 CATCTTGAATGAGATCATTTGATGAAAAACCGTATTAAAGAGGACCGCGAGATTTCCATGAT 180
 |||||
 QY 883 CGATCATCTCTCTCAAACTGCAAGAGACCCAGCCTCAGTACTACACTGACCAAAATCATCAA 942
 |||||
 Db 181 TGATCATCTCTCTCAAACTGCAAGAGACTCAGCCTGAGTACTACACTGACCAAAATCATCAA 240
 |||||
 QY 943 AGGCTTGTCTGTGGCCATGCTTTTTTGGTGAAGTCACTCATCAACTGGGACTTTAGAGTG 1002
 |||||
 Db 241 AGGCTTGTCTGTGGCCATGCTTTTTAGGCGGAACCTGACTCATCAACGGGAACCTTAGAGTG 300
 |||||
 QY 1003 GTCATTATCTAATTTTATGAAATCACCAGAGCTGTGTAAGAGGCAAGAGATGAATGGA 1062
 |||||
 Db 301 GTCGCTATCTAATTTTATGAAATCACCAGAGCTGTGTAAGAGGCAAGAGATGAATGGA 360
 |||||
 QY 1063 CACTCAAGTGGGACAGACCGCTTGTAAATCAGTCAAGACCTTCCAAAACCTTCCATATCT 1122
 |||||
 Db 361 CACTCAAGTGGGACAGACCGCTTGTAAATCAGTCAAGACCTTCCAAAACCTTCCATATCT 420
 |||||
 QY 1123 TAGGAAGATCATCTCTTGAGACACTTAGGTGTGTAACCCCGGCCCAATTTAATACCTCA 1182
 |||||
 Db 421 TAAAAGATCATGCTTTGAGACACTTAAAGTTGTATCTCTCCAGCTCCAATTTCTAATACCTCA 480
 |||||
 QY 1183 TGTGCTTCCAGAAATATTATCAATTTGAGAGATTCAATATCCACGAGACACAAATTGTGAT 1242
 |||||
 Db 481 TGTGCTTCCAGAAATATTATCAATTTGAGAGATTCAATATGTCACGAGACACAAATTGTGAT 540
 |||||
 QY 1243 CATT 1246
 |||||
 Db 541 CATT 544
 |||||

RESULT 12

BQ453555

LOCUS

DEFINITION

BQ453555 564 bp mRNA linear EST 29-MAY-2002
 sa033f08.y1 Gm-cl081 Glycine max cDNA clone SOYBEAN CLONE ID:
 Gm-cl081-2559 5', similar to TR:Q9AFX0 Q9AFX0 CYTOCHROME P450
 MONOOXYGENASE ;, mRNA sequence.

ACCESSION

BQ453555

VERSION

BQ453555.1 GI:21256667

KEYWORDS

EST

Glycine max (soybean)

Glycine max

ORGANISM

Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 564)

Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna

Query Match 22.4%; Score 417; DB 12; Length 425;
Best Local Similarity 98.8%; Pred. No. 4.4e-83;
Matches 420; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1128 AGATCATCTTGTAGACACATTAGGTGTACCCCGCCGCCCAATCTTAATACCTCATGTCT 1187
DB 1 AGATCATCTTGTAGACACATTAGGTGTACCCCGCCGCCCAATCTTAATACCTCATGTCT 60

QY 1188 CTTTCAAGATATTCAATTTGAAGGATTCAATATCCACGAGACACAATTTGATCATTA 1247
DB 61 CTTTCAAGATATTCAATTTGAAGGATTCAATATCCACGAGACACAATTTGATCATTA 120

QY 1248 ATGGTTGGGCATGCGAGAGATCTCAGTTGTGGAAATGATGCCACATGTTTAAACCTG 1307
DB 121 ATGGTTGGGCATGCGAGAGATCTCAGTTGTGGAAATGATGCCACATGTTTAAACCTG 180

QY 1308 AGAGGTTTGTATGGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1367
DB 181 AGAGGTTTGTATGGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240

QY 1368 CTTGCCCGAGAGAACCCATGGCTATGCAAGTGTGAGCTTTTACTTTGGGATTTGATTC 1427
DB 241 CTTGCCCGAGAGAACCCATGGCTATGCAAGTGTGAGCTTTTACTTTGGGATTTGATTC 300

QY 1428 AATGTTTGTGCTGGAACGAGTAAGTGAGGAAAGCTTGATATGACAGAGACAATTTGA 1487
DB 301 AATGTTTGTGCTGGAACGAGTAAGTGAGGAAAGCTTGATATGACAGAGACAATTTGA 360

QY 1488 TCACCTTGTCAAGGTTAATTCATTTGGAGCCATGTCAGGCTCGCCCACTTGCACCTA 1547
DB 361 TCACCTTGTCAAGGTTAATTCATTTGGAGCCATGTCAGGCTCGCCCACTTGCACCTA 420

QY 1548 AATTT 1552
DB 421 AATTT 425

RESULT 15
BI426262
LOCUS
DEFINITION
safo9b06.y3 Gm-c1076 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-c1076-84.5, similar to TF:Q9ZRW6 Q9ZRW6 CYTOCHROME P450 ,
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BI426262 426 bp mRNA linear EST 29-NOV-2001
safo9b06.y3 Gm-c1076 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-c1076-84.5, similar to TF:Q9ZRW6 Q9ZRW6 CYTOCHROME P450 ,
mRNA sequence.

BI426262.1 GI:15203494
EST.
Glycine max (soybean)
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 426)
Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished

TITLE
JOURNAL
COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Putative full length read
vector to vector length is This clone is available through: ResGen,
Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801
For further information call: (800)-533-4363 or contact via email:

ccu@resgen.com
High quality sequence stop: 398.
Location/Qualifiers
1. 426
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1076-84"
/tissue_type="wounded cotyledons"
/dev_stage="11 day old seedlings"
/lab_host="DH10B"
/clone_lib="Gm-c1076"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
from 11 day old seedlings treated with that were treated
with 2 ugs/ml of a crude glucan elicitor preparation
isolated from the mycelial walls of Phytophthora sojae.
The library was prepared using the Stratagene pBluescript
II SK(+) library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with an XhoI restriction site. EcoRI
adaptors were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into E.coli Electromax DH10B host cells. Plant
material was provided by Michael G. Hahn (Complex
Carbohydrate Research Center, University of Georgia) and
the library was constructed by Anu Khanna (Lila Vodkin lab
, University of Illinois)."
BASE COUNT 140 a 58 c 79 g 149 t
ORIGIN

Query Match 22.1%; Score 410.8; DB 12; Length 426;
Best Local Similarity 98.3%; Pred. No. 1.1e-81;
Matches 415; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1408 TACTTTGGGATTTGATTTCAATGTTTGCCTGGAACGAGTAAGTGAGGAAAGCTTGA 1467
DB 5 TTCTGTTTCGAGGGTTGATTTCAATGTTTGCCTGGAACGAGTAAGTGAGGAAAGCTTGA 64

QY 1468 TATGACAGAGACAATTTGGATCACCCTTGTCAAGGTTAATTTCCATTGGAGGCCATGTGCAA 1527
DB 65 TATGACAGAGACAATTTGGATCACCCTTGTCAAGGTTAATTTCCATTGGAGGCCATGTGCAA 124

QY 1528 GGCTCGCCCACTTGCCTCAATAAATTTGAATTTAATTTAATAGTATTTTATTTGGTAA 1587
DB 125 GGCTCGCCCACTTGCCTCAATAAATTTGAATTTAATTTAATAGTATTTTATTTGGTAA 184

QY 1588 ACTTGGGTGATTCAGAACTTAATCTTAATTTTAGTGTGTAAGAGTGGTATCATATA 1647
DB 185 ACTTGGGTGATTCAGAACTTAATCTTAATTTTAGTGTGTAAGAGTGGTATCATATA 244

QY 1648 TACATTTTCAAAATTAATTAATCTTTGTCCAAAATCATCTCATGCAACATATATGCAATT 1707
DB 245 TACATTTTCAAAATTAATTAATCTTTGTCCAAAATCATCTCATGCAACATATATGCAATT 304

QY 1708 GACATCTAGAGAGAAATAGATATAAGAAATTTTATATTTTATTAATCTCTCTTTATCTT 1767
DB 305 GACATCTAGAGAGAAATAGATATAAGAAATTTTATATTTTATTAATCTCTCTTTATCTT 364

QY 1768 ATGTGTCAAGGCCCATTTAGAAATTTGGGTGAGCATTAACATATATCATATATTGTATACCG 1827
DB 365 ATGTGTCAAGGCCCATTTAGAAATTTGGGTGAGCATTAACATATATCATATATTGTATACCG 424

QY 1828 CC 1829
DB 425 CC 426

Search completed: October 21, 2003, 05:47:14
Job time : 3611 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2003, 01:48:19 ; Search time 500 Seconds
(without alignments)
10036.511 Million cell updates/sec

Title: US-09-868-546A-1
Perfect score: 1859
Sequence: 1 gaaacactgacagacga.....aataaattcttacttctc 1859

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1859	100.0	1859	21	AAA29326 Soybean isoflavone
2	1122	60.4	1698	21	AAA29327 Soybean isoflavone
3	396.2	21.3	1497	21	AAC50087 Arabidopsis thalia
4	393.8	21.2	1488	24	ABZ13383 Arabidopsis thalia
5	386.6	20.8	1494	21	AAC43264 Arabidopsis thalia
6	386.6	20.8	1494	24	ABZ14473 Arabidopsis thalia
7	386.6	20.8	1494	25	ABZ42120 Arabidopsis thalia
8	386.6	20.8	1654	21	AAC50097 Arabidopsis thalia

9	384.6	20.7	1655	21	AAC47389 Arabidopsis thalia
10	384.6	20.7	1656	21	AAC37476 Arabidopsis thalia
11	383.8	20.6	1479	24	ABZ14367 Arabidopsis thalia
12	383.8	20.6	1545	21	AAC42545 Arabidopsis thalia
13	383.2	20.6	1503	24	ABZ13668 Arabidopsis thalia
14	381	20.5	1653	21	ABZ14374 Arabidopsis thalia
15	377	20.3	1503	24	ABZ12256 Arabidopsis thalia
16	377	20.3	1664	21	AAC47894 Arabidopsis thalia
17	374.2	20.1	1665	21	AAC50095 Arabidopsis thalia
18	361.8	19.5	1509	24	ABZ13030 Arabidopsis thalia
19	361.8	19.5	1519	21	AAC47053 Arabidopsis thalia
20	354	19.0	1576	21	AAC47416 Arabidopsis thalia
21	354	19.0	1578	21	AAC35968 Arabidopsis thalia
22	351.4	18.9	1674	20	AAAS8406 Jerusalem artichok
23	350.6	18.9	1557	24	ABZ14154 Arabidopsis thalia
24	330.4	17.8	1548	24	ABZ13736 Arabidopsis thalia
25	328.4	17.7	1698	19	AAV63107 Euphorbia lagascae
26	244.4	13.1	843	21	AAA29328 Soybean isoflavone
27	202.8	10.9	722	24	AQE65567 Arabidopsis thalia
28	182.2	9.8	447	21	AA68128 Eucalyptus grandis
29	181.6	9.8	1545	22	AA505178 Taxus cuspidata ox
30	168.6	9.1	475	24	ABL93297 Arabidopsis thalia
31	159.4	8.6	1530	25	AAD47449 Tobacco cytochrome
32	158.2	8.5	1488	24	ABL61191 P. sativum ddfw1 c
33	143.4	7.7	494	21	AAA68129 Eucalyptus grandis
34	139.8	7.5	1770	21	AAA88000 Perilla frutescens
35	139.4	7.5	495	25	ABX61012 Arabidopsis thalia
36	139	7.5	1557	24	ABZ14864 Arabidopsis thalia
37	135.6	7.3	1789	18	AAV94655 Petunia flavonoid
38	131.8	7.1	1704	20	AA60781 Soybean cytochrome
39	130	7.0	1784	21	AAC34631 Arabidopsis thalia
40	129.6	7.0	1770	22	AAF30281 Perilla flavone sy
41	129.4	7.0	1650	21	AAC42321 Arabidopsis thalia
42	128.6	6.9	1812	14	AAQ34489 Insert in pCGP176
43	128.6	6.9	1812	14	AAQ49417 Cytochrome P450 ho
44	128.6	6.9	1824	14	AAQ47842 Flavonoid-3',5'-hy
45	127.2	6.8	522	19	AAV23837 Plant C4H enzyme D

ALIGNMENTS

RESULT 1
AAA29326
ID AAA29326 standard; cDNA; 1859 BP.
XX
AC AAA29326;
XX
DT 26-SEP-2000 (first entry)
XX
DE Soybean isoflavone-2-hydroxylase coding sequence.
XX
KW Soybean; isoflavone-2-hydroxylase; flavonol; biosynthesis; anthocyanin;
KW flower colour; pollen tube; feeding deterrent; UV irradiation; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
CDS 59..1561
FT /*tag= a
FT /product= Isoflavone-2-hydroxylase
XX
PN WO200037652-A2.
XX
PD 29-JUN-2000.
XX
PF 20-DEC-1999; 99WO-US30337.
XX
PR 21-DEC-1998; 98US-0113190.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Famodu OO, McGonigle B, Odell JT, Fader GM, Falco SC;

XX		WPI; 2000-442678/38.
DR		P-P8DB; AAY96593.
PT		New polynucleotide encoding flavonoid biosynthetic enzymes, useful for
XX		producing transgenic plants and immunological screening of cDNA
PT		libraries
XX		Claim 1; Page 30; 36pp; English.
XX		This cDNA, isolated from clone sislc.pk005.n3, encodes a plant (soybean)
CC		isoflavone-2-hydroxylase. It was determined using the sequence of an
CC		isoflavone-2-hydroxylase encoded by a contig composed of clones
CC		s9c1c.pk003.g17, s9e2c.pk004.h7 and slf1.pk0034.g1. The cDNA sequences
CC		can be used for the recombinant production of the enzyme, to isolate
CC		homologues, to create transgenic plants and to provide probes for
CC		genetically and physically mapping genes and as markers for traits linked
CC		to the genes. The proteins can be used for immunological screening, in
CC		particular to raise antibodies against the enzymes. The enzyme and its
CC		gene are useful to study flavonol biosynthesis in plants and provide
CC		means to enhance or otherwise alter flavonol and anthocyanin
CC		biosynthesis. Flavonoids have diverse functions, such as co-pigments in
CC		flower colour, stimulation of pollen tube growth, pollinator attraction,
CC		and feeding deterrents and protection against UV irradiation in fruits
CC		and seeds.
XX		
SQ		Sequence 1859 BP; 536 A; 402 C; 417 G; 504 T; 0 other;
		Query Match 100.0%; Score 1859; DB 21; Length 1859;
		Best Local Similarity 100.0%; Pred. No. 0;
		Matches 1859; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1	GAAACACTGCAGACAGCATAGTCTCTGGTGCAAGAATCAATTGAGCAGCATGGGAAT 60
Db	1	GAAACACTGCAGACAGCATAGTCTCTGGTGCAAGAATCAATTGAGCAGCATGGGAAT 60
QY	61	GTTGTGGTGGTGGTCTCATACGCTGTCTTTTCCTCGGTTCTAATCTCGGCGTGAAGTT 120
Db	61	GTTGTGGTGGTGGTCTCATACGCTGTCTTTTCCTCGGTTCTAATCTCGGCGTGAAGTT 120
QY	121	TGTTTTCAAACAGAAAAATTGAGAAACATACCACGAGTCTCTCTTCCCATAAT 180
Db	121	TGTTTTCAAACAGAAAAATTGAGAAACATACCACGAGTCTCTCTTCCCATAAT 180
QY	181	AGGAACACTTAACCTTCTCGAACGCCAATCCACGGTTCTTCCAACGCAATGTCGAACA 240
Db	181	AGGAACACTTAACCTTCTCGAACGCCAATCCACGGTTCTTCCAACGCAATGTCGAACA 240
QY	241	GTA CGGCAACG TG GTT CCCT CTGGT TC CGG TT CA CGT TC AC GT CT GG CG GTG TCA TCT CCT CTC C 300
Db	241	GTACGGCAACGTGGTTTCCTCTGGTTCGGTTACGCTCTGGCGGTGTGTCATCTCCTCTCC 300
QY	301	AACAGCATACCAAGAAATGCTTCA CAAACA CAG AGTTG C CTTGGC CCA ACCGGCTAC CTT C 360
Db	301	AACAGCATACCAAGAAATGCTTCA CAAACA CAG AGTTG C CTTGGC CCA ACCGGCTAC CTT C 360
QY	361	TCTCTCGGAAAATACATCTTCTCAACAACACACCGTAGGCTTCGTCCTCCACGCGCA 420
Db	361	TCTCTCGGAAAATACATCTTCTCAACAACACACCGTAGGCTTCGTCCTCCACGCGCGCA 420
QY	421	GC ACTGGCC AACCTCCGCGCATCACCGCCCTGGAGCTCCTCTCCACGACGCGCTCCA 480
Db	421	GC ACTGGCGCACTCCGCGCATCACCGCCCTGGAGCTCCTCTCCACGACGCGCTCCA 480
QY	481	CTCCTTCTCCGAAATCCGAGCGGACGAGACGAGCTCTGATGACAGAGTTGGTGTGGC 540
Db	481	CTCCTTCTCCGAAATCCGAGCGGACGAGACGAGCTCTGATGACAGAGTTGGTGTGGC 540
QY	541	CAGAACTCGACGAGGAGAGTTTGC CGGAGTGGAGATAGTTTCAGATGTTCAACGACTT 600
Db	541	CAGAACTCGACGAGGAGAGTTTGC CGGAGTGGAGATAGTTTCAGATGTTCAACGACTT 600
QY	601	AACTTCAACAACATAAATGAGGATGATATCGGGGAAGAGGTTTTTACCGAGGAGGAGTGA 660

Db 1681 TCATCCATGGACAACTATATGTCAATTGACATCTAGAGAGAAATATAGATATATAGATAT 1740
 Qy 1741 TTATATTTTATTACTCTCTTTATCTTATGTGTGTAAGGCCCAATGTAGAAATGGGTGAGC 1800
 Db 1741 TTATATTTTATTACTCTCTTTATCTTATGTGTGTAAGGCCCAATGTAGAAATGGGTGAGC 1800
 Qy 1801 ATTAACATATATCAATATTTGTATACCGCCAGTTTCTCAATATAATTTCTTACTTTC 1859
 Db 1801 ATTAACATATATCAATATTTGTATACCGCCAGTTTCTCAATATAATTTCTTACTTTC 1859

RESULT 2

AA229327

ID AAA9327 standard; cDNA; 1698 BP.

XX AC

XX AAA29327;

XX DT

XX 26-SEP-2000 (first entry)

XX XX

XX Soybean isoflavone-2-hydroxylase coding sequence.

XX DE

XX Soybean; isoflavone-2-hydroxylase; flavonol; biosynthesis; anthocyanin;
 KW flower colour; pollen tube; feeding deterrent; UV irradiation; ss.

XX KW

XX OS

XX Glycine max.

XX XX

XX Key Location/Qualifiers

XX CDS 24..1508

XX FT /*tag= a

XX FT /product= isoflavone-2-hydroxylase

XX FT /partial

XX XX

XX W0200037652-A2.

XX PN

XX XX

XX 29-JUN-2000.

XX PD

XX 20-DEC-1999; 99WO-US30337.

XX XX

XX 21-DEC-1998; 98US-0113190.

XX PR

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX PA

XX Pamodu OO, McGonigle B, Odell JT, Fader GM, Falco SC;

XX PI

XX WPI; 2000-442678/38.

XX DR

XX P-PSDB; AAY96594.

XX XX

XX New polynucleotide encoding flavonoid biosynthetic enzymes, useful for

XX PT producing transgenic plants and immunological screening of cDNA

XX PT libraries

XX XX

XX Claim 1; Page 32-33; 36pp; English.

XX PS

XX This cDNA, isolated from clone arc3c.pk005.f5, encodes a plant (soybean)
 CC isoflavone-2-hydroxylase. It was determined using the sequence of an
 CC isoflavone-2-hydroxylase encoded by a contig composed of clones
 CC sgic1.pk001.g17, sgsc2c.pk004.h7 and slf1.pk0034.g1. The cDNA sequences
 CC can be used for the recombinant production of the enzyme, to isolate
 CC homologues, to create transgenic plants and to provide probes for
 CC genetically and physically mapping genes and as markers for traits linked
 CC to the genes. The proteins can be used for immunological screening, in
 CC particular to raise antibodies against the enzymes. The enzyme and its
 CC gene are useful to study flavonol biosynthesis in plants and provide
 CC means to enhance or otherwise alter flavonol and anthocyanin
 CC biosynthesis. Flavonoids have diverse functions, such as co-pigments in
 CC flower colour, stimulation of pollen tube growth, pollinator attraction,
 CC and feeding deterrents and protection against UV irradiation in fruits
 CC and seeds.

XX SQ

XX Sequence 1698 BP; 495 A; 362 C; 383 G; 458 T; 0 other;

XX XX

XX Query Match 60.4%; Score 1122; DB 21; Length 1698;

XX Best Local Similarity 81.1%; Pred. No. 5.6e-253;

XX Db

Matches 1375; Conservative 0; Mismatches 295; Indels 26; Gaps 5;
 Qy 75 TCTCATACGCTGTCTCTTTCTCTGGTTCATTTCTCTGGCGTGAAGTTCTTTTCCA--AA 131
 Db 25 TATCTTACTCTCTCTTTCTCTGGTTCATTTCTCTGGCGTGAAGTTCTTTTCCAAGAA 84
 Qy 132 GCAGAAATTCAGAAACATATACACAGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 191
 Db 85 GCAGAAATTCAGAAACATATACACAGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 144
 Qy 192 ACCTCTCGAAACAGCAATCCACCGTTTCTTCCAAACGATGTCGAAAACAGTACGCGCAACG 251
 Db 145 ACCTCTCGTGAACACCTATATACACCGTTTCTTCCACCGCATGTCCCAAAAATATGAAACA 204
 Qy 252 TGGTTTCTCTCTGGTTTTCAGCTCTGGCGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 311
 Db 205 TCATATCTCTTGGTTTGGGTTCAGCTCTGGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 264
 Qy 312 AAGAATGCTTCAACCAACACGAGTTGGCTTGGCGCAACGGGTACCTTCTCTCTCTCTCTCT 371
 Db 265 AAGAATGCTTCAACCAACATGATGTACCTTGGCGCAACAGGGGTACGCTCTCTCTCTCTCT 324
 Qy 372 AATACATCTCTTACAAACACACCGTAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 431
 Db 325 AATACATCTCTTACAAACACACCGTAGGGTCTTGTCTCTCTCTCTCTCTCTCTCTCTCT 384
 Qy 432 ACCTCCGCGCATCACCGCT 491
 Db 385 ACCTCCGCGCATCACT 444
 Qy 492 GAATCCGAGCGACGAGACGAGCGTCTGTATGACAGAGTTGGTGTCTGGCCAGAGACTCGA 551
 Db 445 GAATCCGAGCGACGAGACGAGAGGTGTATACACAGGCTGGCCAGGAGCTC----- 496
 Qy 552 ACGAGAGAGAGTTTGGCGAGTGGAGATTAGTTTCAACAGCTTAACTTACAAACA 611
 Db 497 -CGGAAAGATTTCGGCGCGTGGAGATGACCTCCAGTTTGTCTGCTTACGTACAAACA 555
 Qy 612 ACATAATCAGGATGATATCGGGGAAGAGGTTTACGAGAGAGAGTGTAGATGAAGAACG 671
 Db 556 ACATCATGAGGATGATTTTCGGGGAAGCGGTTTACGAGAGAGAGTGAACCTTAAACAG 615
 Qy 672 TTGAGGAGCGAGGAGTTTCAGAGAGCTGTGACAGAAATGTTGAACTCATGGGCTTCG 731
 Db 616 TTGAGGAGCGAGGAGTTTCAGAGACACTGTGAATGAGATGCTGCAACTCATGGGCTTCG 675
 Qy 732 TTAAACAGGAGATCACTTGGCTTCTCTCAGTGGTTTCAATTTTTCAGAAATGTGGAAGC 791
 Db 676 CTAAACAGGAGATCACTTACCTTCTTAAGTGGTTTCAATTTTTCAGAACTGTGGAAGA 735
 Qy 792 GCTTAAAGAGTATCAGTAAGAGGTACGATTCCATCTTGAATGAAGTCTCTTCAATGAGAAC 851
 Db 736 GGTGAAGATATCAGTAAGAGGTATGATACCATCTTGAATGAAGTCTCTTGAATGAGAAC 795
 Qy 852 GTGCCAGCAATGACCGCCAGAGATTCATGATCGATCATCTCTCTCAACTGCAAGAGACCC 911
 Db 796 GTAAACAAAGAGACCGCAGAGATTCATGATTTGGTTCATCTCTCTCAACTGCAAGAGAC 855
 Qy 912 AGCCTCAGTACTACTGACCAAAATCATCAAGGCTTGTCTCTGGCCATGCTTTTGGTG 971
 Db 856 AGCTGCTATTTATACCGCAATCATCAAGGCTTGTCTTGGCTATGCTCTTTGGTG 915
 Qy 972 GAACTGACTCATCAACTGGGACTTTAGAGTGGTCAATTTCTAATTTATTTGAATCAACCCAG 1031
 Db 916 GAAACAGCTCGTCAACTGGAACCTTTAGAGTGGGCATTTCTAATTTAGTGAATGACCCAG 975
 Qy 1032 AGGTGTTGAGAGGCAAGAGATGAATTTGGACACTCAAGTGGGCAAGACCGCTTGTAA 1091
 Db 976 AGGTGCTGCAGAGGCAAGAGATGAGTTGGACCTCAAGTGAAGCAAGATCGGCTGTAA 1035
 Qy 1092 ATGAGTTCAGACCTTCCAAACTTCCATATCTTGAAGAGATCATCTCTTGAACACTTAGGT 1151
 Db 1036 ATGAGTTCAGACCTTCCAAACTTCCATATCTCTGAGAGATAGTCTTGAACACTTAGGT 1095

Qy	1152	TGTACCCCGGCCCCCAATTCATACCTCATGTCTTCAGAAAGATATTACAATTGAAG	1211	PR	29-MAR-1999;	99US-0126785.
Db	1096	TGTAACCTCCGGCTCCCAATTCATATACACACGTTGGCTTCAGAAAGATCAATATCGAAG	1155	PR	01-APR-1999;	99US-0127462.
Qy	1212	GATTCAATATCCCAAGAGACACAAATGTGATCAATTAATGTTGGGATCGAGAGATC	1271	PR	06-APR-1999;	99US-0128234.
Db	1156	GATTCAATGTTCCACGAGACACAATTTGATTAATTAATGTTGGCCATGCMAAGAGATC	1215	PR	08-APR-1999;	99US-0128714.
Qy	1272	CTCAGTTGTGGAATGATGCCACATGCTTTAACTGAGAGGTTTGATGCGAAGGAGAGG	1331	PR	16-APR-1999;	99US-0129845.
Db	1216	CTAAGATATGGAAGATGCGCAAGCTTTAACTGAGAGGTTTGATGGAAGAGGAGAGG	1275	PR	19-APR-1999;	99US-0130077.
Qy	1332	AGAAAAGTTGGTAGCATTTGGCAATGGGAAGAGGGCTTGCCACGAGAGAACCCATGGCTA	1391	PR	21-APR-1999;	99US-0130449.
Db	1276	AGAGAATTTGGTAGCATTTGGTATGGGAAGAGGGCTTGCCACGAGAGAACCCATGGCTA	1335	PR	23-APR-1999;	99US-0130510.
Qy	1392	TGAAAAGTGTGAGCTTTACTTTGGGATTTGATTCATGTTTGACTTGGAAACGAGTAA	1451	PR	28-APR-1999;	99US-0130891.
Db	1336	TGAAAAGTGTGAGCTTTACTTTGGGATTTGATTCATGTTTGACTTGGAAACGAGTAA	1395	PR	28-APR-1999;	99US-0131449.
Qy	1452	GTGAGGAAAGCTTGATATGACAGAGAACCAATTCGATCACCCTTGTCAAGGTTAATTCAT	1511	PR	30-APR-1999;	99US-0132048.
Db	1396	GTGAGAGAACCTTGATATGACAGAGAACCAATTCGATCACCCTTGTCAAGGTTAATTCAT	1455	PR	30-APR-1999;	99US-0132407.
Qy	1512	TGGAGGCCATGTGAAGGCTCGCCCACTTGCCACTTAAATTTGGAATTTAATTTAATAG	1571	PR	04-MAY-1999;	99US-0132484.
Db	1456	TGGAGGCTATGTGAAGCCCGCCCACTCGCCAGCAAGTTGAAGTTAATTAACA	1511	PR	05-MAY-1999;	99US-0132485.
Qy	1572	TATTTTATTTGGTAACTTGGGTGATTGAGATCTAATTAATTAATTTAGTGTGATA	1631	PR	06-MAY-1999;	99US-0132486.
Db	1512	-TATTTTATTTGGTATATTTGGGTG-AGGATCTAATCTCATAATTTGGGTGTGATA	1566	PR	07-MAY-1999;	99US-0132487.
Qy	1632	GAGTGGTGATCATATATACATTTCCAAATTAATTAATTTGTCCTCAAAATCATCCATGA	1691	PR	11-MAY-1999;	99US-0132863.
Db	1567	GTCTATGCAATGTTA-----AAATTAATTAATTTGTCGTATGTCACAGGCCAAATGTA	1621	PR	14-MAY-1999;	99US-0134216.
Qy	1692	CAACTATATGTCAATTCACATCTAGAGAGAAATATAGATATAAGAAATTTATATTTAT	1751	PR	14-MAY-1999;	99US-0134219.
Db	1622	GTACTGGGTGGATTTGCATATACATATCAATATCAATATTGTATAAATCCAGTTTCCTTGAA	1681	PR	14-MAY-1999;	99US-0134221.
Qy	1752	TACTCTCTTTATCTT 1767		PR	18-MAY-1999;	99US-0134768.
Db	1682	TAAATTTCTTTACTTT 1697		PR	19-MAY-1999;	99US-0134941.
RESULT 3				PR	20-MAY-1999;	99US-0135124.
AAC50087				PR	21-MAY-1999;	99US-0135353.
XX				PR	24-MAY-1999;	99US-0135629.
AC				PR	25-MAY-1999;	99US-0136021.
XX				PR	27-MAY-1999;	99US-0136392.
XX				PR	28-MAY-1999;	99US-0136782.
DT				PR	01-JUN-1999;	99US-0137222.
XX				PR	03-JUN-1999;	99US-0137528.
DE				PR	04-JUN-1999;	99US-0137502.
XX				PR	07-JUN-1999;	99US-0137724.
XX				PR	08-JUN-1999;	99US-0138094.
XX				PR	10-JUN-1999;	99US-0138540.
XX				PR	14-JUN-1999;	99US-0139119.
XX				PR	16-JUN-1999;	99US-0139452.
XX				PR	16-JUN-1999;	99US-0139453.
XX				PR	17-JUN-1999;	99US-0139492.
XX				PR	18-JUN-1999;	99US-0139454.
XX				PR	18-JUN-1999;	99US-0139455.
XX				PR	18-JUN-1999;	99US-0139456.
XX				PR	18-JUN-1999;	99US-0139457.
XX				PR	18-JUN-1999;	99US-0139458.
XX				PR	18-JUN-1999;	99US-0139459.
XX				PR	18-JUN-1999;	99US-0139460.
XX				PR	18-JUN-1999;	99US-0139461.
XX				PR	18-JUN-1999;	99US-0139462.
XX				PR	18-JUN-1999;	99US-0139463.
XX				PR	18-JUN-1999;	99US-0139750.
XX				PR	18-JUN-1999;	99US-0139763.
XX				PR	21-JUN-1999;	99US-0139817.
XX				PR	22-JUN-1999;	99US-0139899.
XX				PR	23-JUN-1999;	99US-0140353.
XX				PR	23-JUN-1999;	99US-0140354.
XX				PR	24-JUN-1999;	99US-0140695.
XX				PR	28-JUN-1999;	99US-0140823.
XX				PR	28-JUN-1999;	99US-0140991.
OS				PR	30-JUN-1999;	99US-0141287.
XX				PR	01-JUL-1999;	99US-0141842.
XX				PR	02-JUL-1999;	99US-0142154.
XX				PR	06-JUL-1999;	99US-0142055.
PD				PR	08-JUL-1999;	99US-0142390.
XX				PR	08-JUL-1999;	99US-0142803.
PF				PR	09-JUL-1999;	99US-0142920.
XX				PR	12-JUL-1999;	99US-0142977.
XX				PR	13-JUL-1999;	99US-0143542.
XX				PR	14-JUL-1999;	99US-0143624.
XX				PR	15-JUL-1999;	99US-0144005.
XX				PR	16-JUL-1999;	99US-0144086.
XX				PR	16-JUL-1999;	99US-0144085.

RESULT 3
AAC50087
ID AAC50087 standard; DNA; 1497 BP.
XX
AC AAC50087;
XX
XX 18-OCT-2000 (first entry)
DT
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 63546.
XX
XX Hybridisation assay; Genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.


```
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 21.3%; Score 396.2; DB 21; Length 1497;
Best Local Similarity 57.0%; Pred. No. 7.3e-83;
Matches 838; Conservative 0; Mismatches 603; Indels 30; Gaps 5;

QY 90 TTTTCTCTGTTCTATTCTCTCGCGGGAAGTTTGTTCCTCCAAAGCAGAAATTTGAGAAACA 149
DB 29 TTCTCTTCTGTTGTTCTCTCTCTCATTTACTTAATTTGAAATCTCAAGCGAAGCCAAATC 88
QY 150 TACCACCAAGTCTCTCC---TCCTCTTCCCATTAATAGGAAACCTTAACCTCTCGAAGCAGC 206
DB 89 TACCTCCGAGTCCGGCATGGTCATTACCGGTGATTGGTCATCTCGCCCTTCTCAAAACCAC 148
QY 207 CAATCCACGGTTTCTTCCAAACGATGTCGAACAGTACGGCAACG-----TGGTTTCCC 260
DB 149 CGAATTCATCGCACATTTCTCTCCCTCTCTCAATCCCTAAACAATGCTCCGATCTTCTCCC 208
QY 261 TCTGTTTCGTTTCAGTCTGCGCGTTGTCTCTCTCTCCAAACAGCATACCAAGAAATGCT 320
DB 209 TCCGACTCGTAAACCGACTCGTTTTTCGTCAACTCTGCACACTCGATCGCGGAGGATGTT 268
QY 321 TCACCAAAACAGACGCTTGGCCAAACCGGCTACCTTCTCTCTCGGAAAAATACATCT 380
DB 269 TCACCAAAACAGACGCTGCTACTGGCGAACAGACCAAACTTCATCTCTCGCTAAACACGTTG 328
QY 381 TCTACAAACAGACCGCTAGGCTCTGCTCCCGGGGAGCAGCTGGCGCAACCTCCGCC 440
DB 329 CGTACGATTACAAACCATGATCGCAGCTTCTTACGGTGACCACTGGCGGTAACTCCGCC 388
QY 441 GCATCACCCTCGGACGCTCTCTCCAGCAGCGGTCACCTCTCTCTCCGGAATCCGGA 500
DB 389 GCATCGGCTCCGTCGAGATATTCTCCAAATCACCGTCTCAATAGCTTTTGTCTATTCTGTA 448
QY 501 GCGACGAGCAAGCGTCTGATGCGAGAGTTGGTGTGCGCCCAAGAACTCGAAGCGAAGAG 560
DB 449 AAGACGAGATCCGACGACTTGGTTTCGTTCTCTCGGAAC-----TTTTCACAAG 499
QY 561 AGTTTCCGCGAGTGGAGATTAGTTTCGATGTTTCAAACGACTTAATTACACACATTAATCA 620
DB 500 AGTTTGTGAAAGTGGATATGAATCAATGTTTATCTGACTTAACATTTCAACAACTTTTAA 559
QY 621 GGATCATATCGGGAAGAGGTTTTCACGGAGAGGAGGTGAGATGAAGAACGTTTGAGGAAG 680
DB 560 GAATGTTGGCCGAAACGTTTACTACGGAGACGTTGTGAGG-----ATGATCCGGAGG 613
```


Db 467 GGCTCATACACATCTCTAGAGACT-----CCTTGACGGATTGTTGAAGTGG 517
Qy 576 AGATTAGTTGATGTTCAACGACTTAACCTTACAAACATATATGAGATCATATCGGGA 635
Db 518 AGTTGAAATCGTTATTAAACCACTTGGCATTCACCAACATCATCATGATGGTAGCGGAA 577
Qy 636 AGAGGTTTTACGAGAGGAGAGTGAGATGAAGAACGTTGAGGAGCGAGGAGTTGAGAG 695
Db 578 AAGCATATTACGG-----TACCGGTACAGAGCAATGATGAAGCAAGCAAGCTCGTAGGG 631
Qy 696 AGACTGTGACAGAAATGTTGGAATCATGGCTTGTCTAACAGGAGAGATCACTCCCTT 755
Db 632 AACTTATAGCGGAGATAATGGCCGGCGCGCTTCTGGAATTTGGCTGATATATCTCCGT 691
Qy 756 TCCTCAGTGGTTTCGATTTTTCAGAAATGTGAGAGCGCTTAAAGATATCAGTAAGAGT 815
Db 692 CCAATAATTGGGTC-----ACAACTTTGAGAACAGACCAAAATCTTTGGGGAATCGAC 745
Qy 816 ACGATTCCATCTTGAATAGATCCCTTCATGAGAACCGTGCAGCAATGACCGCCAGATT 875
Db 746 TCGATAGAGTCTTGCAGAAATCTGGTTGATGAGAAACGTCGAGAGAAAGAAAGGGTCAAA 805
Qy 876 CCATGATCGATCATCTCCTCAAACTGCAAGAGACCCAGCTCAGTACTACACTGACCAA 935
Db 806 CTTTGATCGATCATTTGCTTCTTTCAGAGAAACCGAACCCGAGTACTATCTGATGTCA 865
Qy 936 TCATCAAGGCTTGTCTGCGCATGCTTTTGGTGAACCTGACTCATCACTGGGACTT 995
Db 866 TCATCAAGGAAATCATACTGGCTTGTCTTGGGGGACAGATACATCTTCAGTGAAGT 925
Qy 996 TAGAGTGGTCATTATCTAATTTATGTAATCACCAGAGAGTGTGAAGAGGCGCAAGAGTG 1055
Db 926 TGAATGGGCAATGTCAAAATTTGCTGAACCATCCAGAGATACCTTTGAGAAAGCGAGAGCG 985
Qy 1056 AATTGGACACTCAAGTGGGCAAGACCGCTTGTAAATGAGTCAAGCTTCCAAACTTC 1115
Db 986 AGATCGATGATAAATCGGTTGACCGGTTAGTGAAGAAATCAGATATTTGAATCTCC 1045
Qy 1116 CATATCTTAGGAAGATCATCTTTGAGACACTTAGTGTGACCCCGGCCCAATTTCTAA 1175
Db 1046 ATTATCTCAAAACATTTGTGCAAAACATTACGTTTGTATCTCGGTTCCACTACTAC 1105
Qy 1176 TACCTCATGTTCTTCAAGATATTACAAATTTGAAGATTCATATCCACGAGACAA 1235
Db 1106 TCCTCATTTCTCATCGGATGAATGTAAAGTGGCGGCTACGATATGCCACGTCGCACGT 1165
Qy 1236 TTGTGATCATTAATGTTGGGGCATGCAGAGAGATCTCTCAGTTGTGGAATGATGCCACAT 1295
Db 1166 TGTATTAAACAGTATGGCGCATGCATAGATCCAGTTTTATGGAGAGCCAGAGA 1225
Qy 1296 GCTTTAAACCTGAGAGTTGATGTGGAAGGAGAGAGAAAGTTGGTAGCAATTTGGCA 1355
Db 1226 GGTTCGAAGCGGAGAGGTTTCAGAGAAAGAGAGAGGCTTCGAAGCTAATGCCGTTTGGGA 1285
Qy 1356 TGGGAGAGGCGCTGCCAGGAGAACCCATGCTATGCAAACTGTCAAGCTTACCTTTGG 1415
Db 1286 TGGGAGACAGCTTGTCTGAGAGCTGAGCTTGGGAGCGGTTAGTGAGCCTTGTCTCTT 1345
Qy 1416 GATTGTTGATTAATGTTTTCAGTGGAAACGAGTAAGTGAAGAAAGCTTGAATGACAG 1475
Db 1346 GGTGCTTGTATTCAGTCTTTCAGATGGGAGAGAGTTGGTGAGAACTTGTGGACATGACTG 1405
Qy 1476 AGAACAAATGGATCAGCTTCTCAAGGTTAATTCATTTGAGGCGCATGTGCAAGGCTGCC 1535
Db 1406 AAGGCGAAGGATCAGTATGCTTAAAGCTACTCCGTTGCGAGCTATGTGCAAGGCGCAGTG 1465
Qy 1536 CACTTGGCACTAAA 1550
Db 1466 CCATTGTTGGTAAA 1480

RESULT 5
AAC43264

ID AAC43264 standard; DNA; 1494 BP.
XX
AC AAC43264;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 38633.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.

Db 209 TCCGACTCGGTAAACCGACTCGTTTTCGTGAACCTGTCACACTCGATCGCGGAGAAATGTT 268
Qy 321 TCACCAAAACAGCGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 380
Db 269 TCACCAAAACAGCGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 328
Qy 381 TCTACCAACACACCGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 440
Db 329 CGTACGATTAACAAACCGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 388
Qy 441 GCATCAACCGCTTCGAGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 500
Db 389 GCATCGCTTCGAGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 448
Qy 501 GCGACGAGACGAGTTCGAGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 560
Db 449 AAGACGAGATCCGACGACTTGTGTTCTCTCTCGGGAATAATCATCT 499
Qy 561 AGTTTGGCGAGTTCGAGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 620
Db 500 AGTTTGGCGAGTTCGAGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 559
Qy 621 GGTATGATATCGGGAAGAGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 680
Db 560 GAATGGTGGCGGAAGAGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 613
Qy 681 CGAGGAGTTCAGAGACTGTGACAGAAATGTTGGAACCTCATGGGCTTGGCTTAAACAAGG 740
Db 614 CTAACGCTGCGGAGCTTATAGCGGATGTGTGCTTGTGCTGTTGGAACCGCTG 673
Qy 741 GAGATCATCTTCTCTCGGGAAGAGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 800
Db 674 TTGATTAATCTACCGGTTTTCGCGTTCG-----TTTCAGATTACGAGACACGCGGTAAAGA 727
Qy 801 GTATCAGTAGAGTACGATTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 860
Db 728 AGTTTGGCGGAGTTCGAGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 787
Qy 861 ATCAGCCGCAAGATTCAGTTCCTTGGCCAAACCGGTACCTTCTCTCTCGGGAATAATCATCT 920
Db 788 AGGAGAAAGAAACACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 847
Qy 921 ACTACAGTACCAATCATCAAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 980
Db 848 ACTTACGAGTATCATCAAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 907
Qy 981 CATCAACTGGGACTTTAGAGTGGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1040
Db 908 CATCAGGGTTACGTAGATGGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 967
Qy 1041 AGAAGCAAGAGATGAATTCGACACTCAAGTGGGCAAGACCGCTTGTAAATGAGTTCAG 1100
Db 968 ACAAGCGAGATGAATTCGACACTCAAGTGGGCAAGACCGCTTGTAAATGAGTTCAG 1027
Qy 1101 ACCTTCCAAACTTCATATCTTAGGAGATCATCTTGGAGACACTTAGTGTGACCCCC 1160
Db 1028 ATATCTCAAACTTCATATCTTAGGAGATCATCTTGGAGACACTTAGTGTGACCCCC 1087
Qy 1161 CGGCCCCAATTTAAATACCTCATGTGCTTCAAGAGATATTACAAATTAAGAGGATTCAAATA 1220
Db 1088 CGGCTCCATGCTTCT 1147
Qy 1221 TCCACGAGACAAATTTGATCATTAATGTTTGGGCGATGCGAGAGATTCCTCAGTTGT 1280
Db 1148 TGCCGGTGGCAGGATTAATGACCAATGTGTGGGCTATACACAGAGATTCCTCAGTTGT 1207
Qy 1281 GGAATGATCCACATGCTTTAACTGAGAGTTCGATGAGAGAGGAGAGAGAGAGAGAGAGAG 1340
Db 1208 GGGATGATCCAAATGAGCTTCAAGCCGAGAGAGGTTTGAGAAAGAGAGAGAGAGAGAGAG 1267
Qy 1341 TGTGTAGCATTTGGCATGGGAGAGGCTTGGCCAGGAGAGACCATGCTGCTATGCAAAAGT 1400
Db 1268 TAATGCCGTTTGGTTAGGAGAGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1327

Qy 1401 TCAGCTTTACCTTTGGGATTTGATTCATGATTTTGAATGATTTGACTGGAAACGAGTAAATGAGGAAA 1460
Db 1328 TAAACCTGACTCTTGGATCAITGATTCAGTGTGTTGAATGGGA---GAAGATTGGAGAAG 1384
Qy 1461 AGCTTGTATGACAGAGAACATTTGGATGATCACCCTGTCAAGTTAATTCATTGGAGGCCA 1520
Db 1385 AAGTGGATATGATGAGGCAAGGTGTACAAATGCTTAAAGCCAGCCTTTTGGAAAGCCA 1444
Qy 1521 TGTGCAAGCTGCCCACTTGCACCTAAAT 1551
Db 1445 TGTGACAGACGCTCCCTCTGTTGTTAAAT 1475

RESULT 6

ABZ14473
ID ABZ14473 standard; DNA; 1494 BP.

XX ABZ14473;

XX DT 21-JAN-2003 (first entry)

XX Arabidopsis thaliana stress regulated gene SEQ ID NO 2278.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.

XX WO200216655-A2.

XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US26685.

XX 24-AUG-2000; 2000US-227866P.

XX 26-JAN-2001; 2001US-26447P.

XX 22-JUN-2001; 2001US-30011P.

XX (SCRI) SCRIPPS RES INST.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

XX Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses

XX Claim 144; SEQ ID NO 2278; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising:

XX (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and

XX (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the

XX production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that

XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used in methods of the invention.

XX Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by

XX the European Patent Office.

XX SQ Sequence 1494 BP; 399 A; 341 C; 355 G; 399 T; 0 other;

Query Match 20.8%; Score 386.6; DB 24; Length 1494;

Best Local Similarity 57.0%; Pred. No. 1.3e-80;

Matches 839; Conservative 0; Mismatches 599; Indels 33; Gaps 6;

Qy 90 TTTTCTGTTCTATCTCTCGGCGTGAAGTTGTTTTCCTCAAGCAGAAAATTGAGAAACA 149

PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0139454.
PR 14-JUN-1999; 99US-0139847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139847.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148665.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 08-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.

	14-OCT-1999;	99US-0159329.	Query Match	20.8%;	Score 386.6;	DB 21;	Length 1654;
	PR 14-OCT-1999;	99US-0159330.	Best Local Similarity	57.0%;	Pred. No. 1.3e-80;		
	PR 14-OCT-1999;	99US-0159331.	Matches 83;	Conservative 0;	Mismatches 599;	Indels 33;	Gaps 6
Qy	90	TTTTCTCGTGTCTATTCCTCGCGTGAAGTTGTTTTCCTCCAAAGCAGAGAAAATTTGAGAAACA	149				
Db	46	TTCTCTCGTGTGTTCTCTCTCTCTCATTTACTTAATTTGAGAAACTCAAGCGAAAGCCCAATC	105				
Qy	150	TACACACAGGTCCTCC---TCTCTTCCATAATAGGAAAACCTTAACTCTCTCGAACAGC	206				
Db	106	TACCTCGAGTCGCGCATGTCATTAACCGGTGATTTGTCATCTCCGCCCTTCTCAAACCCAC	165				
Qy	207	CAATCCACCGTTCTTCCACAGCATGTCGAAACAGTACGGCAAGC-----TGGTTTCCC	260				
Db	166	CGATTATCGACATTCCTCTCCCTCTCTCAATCCCTAAACAATGCTCGATCTTCTCTCCC	225				
Qy	261	TCTGGTTTCGGTTCACGTCGTGGCCGTTGTTCATCTCTCTCCACAGCATACCAAGAATGCT	320				
Db	226	TCCGACTCGGTAAACCGACTCGTCTTTTCGTGAACTCGTCACACTCGATCGCGAGGAATGTT	285				
Qy	321	TCACAAAACAGAGTTCGCTTGGCCAAACCGGCTACCTTCTCTCTCGGGAATAATCATCT	380				
Db	286	TCACAAAACAGCGTCGTACTTGGCGAAACAGACCAAACCTTCATCTCGCTAAACACGTG	345				
Qy	381	TCTACAAACACACACCGTTCGCTCTCTCCACGGCGAGCAGTTCGCGCACTCCGCC	440				
Db	346	CGTAGGATTACAAACCAATGATCGCAGCTTCTACGGTGAACACHTGGGTAACTCCGCC	405				
Qy	441	GCATCACGCCCTGGAAGTCTCTCTCCACGACGCGTCCACTCTTCTTCGGAATCCGGA	500				
Db	406	GCATCGGCTCCGTTCGAGATATTTCTCCAATCACCGTCTCAATAGCTTTTTTGTCTATTCTGA	465				
Qy	501	CGCAGCAGACGAAGCGTCTGATGACAGAGTTGGTCTCGGCCAAGAACTCGAACGAGGAAG	560				
Db	466	AAGACGAGATCCGACGACTTGTGTTTCGTTCTCTCGGAAC-----TTTTCCACAAG	516				
Qy	561	AGTTTTCGCGAGTGAGATTAAGTTTCGATGTTTCAACGACTTAATCTTACAAACAATAATCA	620				
Db	517	AGTTTGTGAAAGTGGATATGAATCAATGTTATCTGACTTTAACTTCAACAACATTTTAA	576				
Qy	621	GGATCATATCCGGGAAGAGGTTTTCACGAGAGGAGATGACATGAAGAACGTTTGAGGAAG	680				
Db	577	GAATGGTGGCCGAAAAACGTTACTACGGAGACGGTGTG-----AGGATGATCCGAGG	630				
Qy	681	CGAGGAGTTCAGAGAGACTGTGACAGAAAATGTTGAAACTCATNGGCTTGGCTTAACAAG	740				
Db	631	CTAAACGTTGTCGGCAGCGCTTATAGCGGATGTGGTGGCTTGTCTGGTGTCTGGAACGCTG	690				

XX PF1033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126284.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149502.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.

QY 1564 A 1564
Db 1565 A 1565

RESULT 10

AAC37476
ID AAC37476 standard; DNA; 1656 BP.

XX AC AAC37476;
XX AC

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 17517.

XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.
XX AC

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136332.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140623.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.	Query Match	Score 383.8; DB 21; Length 1545;
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.	Best Local Similarity	57.2%; Pred. No. 5.9e-80;
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.	Matches 794; Conservative	0; Mismatches 557; Indels 36; Gaps 4;
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.		
PR	22-JUL-1999;	99US-0145145.	PR	21-OCT-1999;	99US-0160980.		
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.		
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.		
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.		
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.		
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.		
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.		
PR	28-JUL-1999;	99US-0145921.	PR	26-OCT-1999;	99US-0161360.		
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.		
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.		
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.		
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.		
PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0161993.		
PR	04-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.		
PR	05-AUG-1999;	99US-0147192.					
PR	05-AUG-1999;	99US-0147260.					
PR	06-AUG-1999;	99US-0147303.					
PR	06-AUG-1999;	99US-0147416.					
PR	09-AUG-1999;	99US-0147493.					
PR	09-AUG-1999;	99US-0147935.					
PR	10-AUG-1999;	99US-0148171.					
PR	11-AUG-1999;	99US-0148319.					
PR	12-AUG-1999;	99US-0148341.					
PR	13-AUG-1999;	99US-0148565.					
PR	13-AUG-1999;	99US-0148684.					
PR	16-AUG-1999;	99US-0149368.					
PR	17-AUG-1999;	99US-0149175.					
PR	18-AUG-1999;	99US-0149426.					
PR	20-AUG-1999;	99US-0149722.					
PR	20-AUG-1999;	99US-0149723.					
PR	20-AUG-1999;	99US-0149929.					
PR	23-AUG-1999;	99US-0149902.					
PR	23-AUG-1999;	99US-0149930.					
PR	25-AUG-1999;	99US-0150566.					
PR	26-AUG-1999;	99US-0150884.					
PR	27-AUG-1999;	99US-0151065.					
PR	27-AUG-1999;	99US-0151066.					
PR	27-AUG-1999;	99US-0151080.					
PR	30-AUG-1999;	99US-0151303.					
PR	31-AUG-1999;	99US-0151438.					
PR	01-SEP-1999;	99US-0151930.					
PR	07-SEP-1999;	99US-0152363.					
PR	10-SEP-1999;	99US-0153070.					
PR	13-SEP-1999;	99US-0153758.					
PR	15-SEP-1999;	99US-0154018.					
PR	16-SEP-1999;	99US-0154039.					
PR	20-SEP-1999;	99US-0154779.					
PR	22-SEP-1999;	99US-0155139.					
PR	23-SEP-1999;	99US-0155486.					
PR	24-SEP-1999;	99US-0155659.					
PR	28-SEP-1999;	99US-0156458.					
PR	29-SEP-1999;	99US-0156596.					
PR	04-OCT-1999;	99US-0157117.					
PR	05-OCT-1999;	99US-0157753.					
PR	06-OCT-1999;	99US-0157865.					
PR	07-OCT-1999;	99US-0158029.					
PR	08-OCT-1999;	99US-0158232.					
PR	12-OCT-1999;	99US-0158369.					
PR	13-OCT-1999;	99US-0159293.					
PR	13-OCT-1999;	99US-0159294.					
PR	13-OCT-1999;	99US-0159295.					
PR	14-OCT-1999;	99US-0159323.					
PR	14-OCT-1999;	99US-0159330.					
PR	14-OCT-1999;	99US-0159331.					
PR	14-OCT-1999;	99US-0159637.					
PR	14-OCT-1999;	99US-0159638.					
PR	18-OCT-1999;	99US-0159584.					
PR	21-OCT-1999;	99US-0160741.					
PR	21-OCT-1999;	99US-0160767.					

QY 945 GCCTTCTCTGGCCATGCTTTTGGTGGAACTGACTCATCACTGGGACTTTAGAGTGT 1004
DB 932 GAATCATACTTACCTTGATAAATGGCAGGACAGATACGTCATCAGTAACACTAGAAATGG 991
QY 1005 CATTATCTAATTTATTGAATCATCCAGAGGTGTTGAAGAAGCAAGAGATGAATGGACA 1064
DB 992 CAATGTCAAATCTGTGTGAACCATCCAGAAATATCTTAAGAAAGCGAGATGGAAATCGATG 1051
QY 1065 CTCAAGTGGGACAAAGCCGCTGTGTTAAATGAGTCAGACCTTCCAAAATTTCCATATCTTA 1124
DB 1052 AAAAGTGGTGTAGACCGATTAGTAGACGAATCGGACATTTGTAATCTCTCTTATCTCC 1111
QY 1125 GGAAGATCATCTTTGAGACATTTAGGTGTACCCCGCCGCCAATTTCTAATACCTCATG 1184
DB 1112 AAAGCATTTGTTGGAACACATACGTATGTACCCGCGAGTCCCACTACTACTACCTCAT 1171
QY 1185 TGTCTTCAAGAATATTACAATTTGAAGGATTCAATATCCACGAGACACAATTTGTGATCA 1244
DB 1172 TGTATCAGAAGATTGTAAGTTGGAGGCTACGATATCAAGTGGAAACATTTGTAATGA 1231
QY 1245 TTAATGTTGGGCGATCGAGAGATCCTCAGTTGTGGAATGATGCCATGCTTTAAAC 1304
DB 1232 CCAACCATGGGCGCATAGATAGATCCAGAGTATGGGAAGATCCTGAGATATTCAAAC 1291
QY 1305 CTGAGAGTTTGTATGTGGAAGAGAGAGAGAAAGTTGGTAGCATTTGGCATGGGAGAA 1364
DB 1292 CAGAAAGATTGGAAGAGAGAGAGAGCTGAGAAGCTAATCTCATTTGGGATGGGACGAA 1351
QY 1365 GGGCTTGGCCAGGAGAACCATGCTATGCAAGTGTGACGTTTACTTTGGGATTTGTCGA 1424
DB 1352 GAGCTTGTCTGGGCGGCGTAGCTCATCGGCTAATAACCAAGGCTCTTTGGAAGTTTGG 1411
QY 1425 TTCAATGTTTGAAGTGAACAGAGTAAGTGAAGGAAAGCTTGAATGACAGAGAAACAAT 1484
DB 1412 TTCAATGTTTGAAGTGAAGGAAAGTGTGAGGATTTTGGACATGACCGAGACAAG 1471
QY 1485 GGAATCACTTGTCAAGGTTAATTTCCATTTGAGGCGCATGTGCAAGGCTCGGCCACTTGCCA 1544
DB 1472 GAGCCACATTTGCCAAAGCTATACCATTAAGAGCCATGTGCAAGGACGCTTCTATTGTTG 1531
QY 1545 CTAAAT 1551
DB 1532 ATAACT 1538

RESULT 13

ABZ13668
ID ABZ13668 standard; DNA; 1503 BP.

AC ABZ13668;

DT 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1473.

KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

OS Arabidopsis thaliana.

PN WO200216555-A2.

PD 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US26685.

XX 24-AUG-2000; 2000US-227866P.

PR 26-JAN-2001; 2001US-264647P.

PR 22-JUN-2001; 2001US-300111P.

XX (SCRI) SCRIPPS RES INST.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX

PI Harper JF, Krepe J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

XX Identifying a stress condition to which a plant cell has been exposed
PT and producing plants with increased tolerance to these abiotic stresses

PS Claim 144; SEQ ID NO 1473; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising:
CC (a) contacting nucleic acid representative of expressed polynucleotides
CC in the plant cell with an array or probes representative of the plant
CC cell genome; and
CC (b) detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention.

CC Note: The sequence data for this patent is not represented in the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.

XX Sequence 1503 BP; 429 A; 391 C; 322 G; 361 T; 0 other;

Query Match 20.6%; Score 383.2; DB 24; Length 1503;

Best Local Similarity 56.5%; Pred. No. 8.1e-80;

Matches 847; Conservative 0; Mismatches 618; Indels 33; Gaps 6;

QY 85 TGTCTTTTCTGTTTCTATTCCTCGGCGTGAAGTTTGTTCCTCAAGCAAGAAATGAG 144

DB 18 TCTCCTCTTCTATTTTGTAAAGTTACAAAATTTCTACTCCAAAACGCAACGTTT 77

QY 145 AAACATACACAGAGTCTCTCTCTTCCATATAGGAAACCTTAACCTCTCGAAACA 204

DB 78 CAACCTTCTCCAGAGACCGCTCAGACCTTGTGTCGCCATCTCCACCTCATGAAACC 137

QY 205 GCCAATCCACCGTTTCTTCCAAACGATGTGAAACAGTACGGCAACGTGGTTTCCCTCTG 264

DB 138 GCGGATCCACCGTCTCTCTCCAAACGCTACTCTCAACCAATACGCGCCAAATCTTTTCCCTCG 197

QY 265 GTTGGTTACGTCGTGGCGGTTGTCTCTCTCCACAGCATACCAAGATGTTCTAC 324

DB 198 TTTCCGCTCCCGCGCTGCGGTGATCAATCAATCCCTTCTCTCGCCCAAGAAATCTCTCAC 257

QY 325 ---CAAAACAGAGTTGCTTGGCCAAACCGGTACTCTCTCTCGGAAATAATACATCTT 381

DB 258 CGGCCAAAACGACATCGTTCTCTCCAGCGCGCGCTCCAACTCACCGCCAAATAGTCGC 317

QY 382 CTACAACAAACACCGTAGGCTCTGTCTCCACGCGAGCATCTGGCGCAACCTCGCGCG 441

DB 318 ATACAACCAACACACCGTTGGTACAGCTCTTACGCGGACCACTGGCGTAACCTCGCGCG 377

QY 442 CATCACCGGCTTGGAGCTCTCTCCAGCGCGGTCCACTCTCTCTCGGATCCGAG 501

DB 378 CATTTGCTCCCAAGAGATTTCTCTCTCACCGCTCTCATCACTTCCAAACATCCGCAA 437

QY 502 CGACGAGACGAAGCGCTGTGATGACAGAGTTGTGTGCGCCAAAGAACTCGAAACGAGAA-- 559

DB 438 AGACGAGATCTTACGGATGCTCACAGGCTCTCACGTTACACACAAACCTCCACGAAAG 497

QY 560 ----CAGTTTGGCGAGTGGAGATTAGTTTGGATTTCAACGCTTAACTTACAAACAT 615

DB 498 CAACGATTTCACTCATATCGAGCTAGAACCCTCTTATCTGATCTAAACATTCACAAACAT 557

QY 616 AATGAGGATGATATCGGGAAGAGGTTTTCAGGAGGAGAGTGAATCAAGAACTTGA 675

DB 558 AGTAAGAATGGTTACAGGGAAGAGATATTACGCGCACGACG-----TCAACAAACAGA 611

QY 676 GGAACGAGGAGTTTCAGAGAGACTGTGACAGAAATGTTGGAATCTCATGGGCTTGGCTAA 735

Db	612	AGAAGCAGAGCTATTCAAGAAGCTAGTCTACGATATGCCATGTATAGCGGCGCTAATCA	671	OS	Arabidopsis thaliana.
Qy	736	CAAGGGAGATCACTTCGCTTCTCAGGTGGTTCGATTTTCAGAAATGTGGAAGCGTT	795	XX	EP1033405-A2.
Db	672	TTCCGCTGATTAATACCAATATCAAACTATTCGGAA---ACAAATTCGAGAAAGAGT	728	PD	06-SEP-2000.
Qy	796	AAGAGTATCAGTAAGAGGTACGATCCATCTTTGAATAAGATCCTTCATGAGAACCGTGC	855	XX	25-FEB-2000; 2000EP-0301439.
Db	729	TAAAGCTATAGGCAAAATCAATGATGATATTTTGCAGCGTTTGTAGATGAAT--GTAG	785	XX	25-FEB-1999; 99US-0121825.
Qy	856	CAGCAATGACCGCCAGATTCATGATCGATCATCTCTCAAACTCGAAGACCCAGCC	915	PR	05-MAR-1999; 99US-0123180.
Db	786	AAGAGATAAAGAGGTAACAATGTCATATTTGATCTCTCTACACAACAACAAC	845	PR	09-MAR-1999; 99US-0123548.
Qy	916	TCAGTACTACACTGACCAAAATCATCAAGGCCCTTCTCGGCCATGCTTTTGGTGGAA	975	PR	23-MAR-1999; 99US-0125788.
Db	846	TGAGTATTACACTGACGTTATCATCAAGGCCCTTATGATGTCATGCTCGCCGAA	905	PR	25-MAR-1999; 99US-0126264.
Qy	976	TGACTCATCACTGGGACCTTTAGAGTGGTCATTAATCTAATTTAATGAATCACCGAGGT	1035	PR	29-MAR-1999; 99US-0126785.
Db	906	TGAGACCTCGCTGTGACATTAGAGTGGCGATGGCGAATCTACTGAGAAATCCAGAA	965	PR	01-APR-1999; 99US-0127462.
Qy	1036	GTTGAAGAGGCAAGAGATGAATGGACACTCAAGTGGACACAGCGCTTGTTAATCA	1095	PR	06-APR-1999; 99US-0128234.
Db	966	ACTGGAGAGGCGAGATCGGAGATCGATGAGAAGATCGGAAAGATCGACTAATCGAC	1025	PR	08-APR-1999; 99US-0128714.
Qy	1096	GTGAGACCTTCCAAAATCCATATCTTAGAAGATCATCTTTGAGACACTTAGGTTGTA	1155	PR	16-APR-1999; 99US-0129845.
Db	1026	ATCAGACATCGAGTCTCTTACCTCCAAAACGTCGTCGGAACATTCGGGTATT	1085	PR	19-APR-1999; 99US-0130077.
Qy	1156	CCCCCGGCCCCAATTCATACTCATGCTTTCAGAAAGATTAATCAATTTGAAGAGAT	1215	PR	21-APR-1999; 99US-0130449.
Db	1086	TCCGGTGGCGCGCTTTCTAATCCCTCGATCACCGAGGACGACATGAAAATCGAGG	1145	PR	23-APR-1999; 99US-0130510.
Qy	1216	CRATATCCACGAGACACATTTGATCATTAATGTTGGGGATCGACGAGATCCTCA	1275	PR	28-APR-1999; 99US-0131449.
Db	1146	CGAGTCCCGGAGATACGATGATGATGTTAAGCTTGGGCTATACATAGATCCAGA	1205	PR	30-APR-1999; 99US-0132048.
Qy	1276	GTTGTGGAATGATGCCACATGCTTTAAACCTGAGAGGTTTGTGGAAGGAG------	1328	PR	30-APR-1999; 99US-0132484.
Db	1206	GATTTGGAGAGCCAGAGAAGTTTAAATCGATCGATATAATGACGGATCGGAGCGA	1265	PR	04-MAY-1999; 99US-0132484.
Qy	1329	-----AGGAGAAAAGTTGTAGCATTTGGCATGGGAGAGGCTTCCAGGAGAAC	1383	PR	05-MAY-1999; 99US-0132485.
Db	1266	TTACTATGTTTACAAGCTGATGCGTTTGGGAATGGCCGAGAACCTTGTCCCGCGCC	1325	PR	06-MAY-1999; 99US-0132486.
Qy	1384	CATGGCTATCAAAAGTCTCAGCTTTACTTTGGGATTTGATTCATGTTTTCACCTGGA	1443	PR	06-MAY-1999; 99US-0132487.
Db	1326	ATTAGGTCAGAGATTGTGACATTGGCGCTTGGATCGTTGATTCATGCTTTGAATGG	1385	PR	07-MAY-1999; 99US-0132863.
Qy	1444	ACGAGTAAGTGAGGAAAAGCTTGATATGACAGAGAACAAATTTGGATCACCTTGTCA	1503	PR	11-MAY-1999; 99US-0134256.
Db	1386	GAATGTGAAGGGAAGAGATGATATGCTGAGAGTACTGGCTTGGGTATGCTTAAG	1445	PR	14-MAY-1999; 99US-0134218.
Qy	1504	AATTCCATTGGAGCCCATGTCAGGCTCGCCCATTTGCCCATTAATAATTTGAATTTAA	1561	PR	14-MAY-1999; 99US-0134221.
Db	1446	GGATCCTTTACGGGCCATGTGTAGGCTTAGGCCCATTAATGTCAAAGCTTTTACTTAA	1503	PR	18-MAY-1999; 99US-0134768.
RESULT 14					99US-0134768.
AAC41374					99US-0134768.
ID	AAC41374 standard; DNA; 1653 BP.				99US-0134768.
XX					99US-0134768.
AC					99US-0134768.
XX					99US-0134768.
DT	17-OCT-2000 (first entry)				99US-0134768.
XX					99US-0134768.
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 31648.				99US-0134768.
XX					99US-0134768.
KW	Hybridisation assay; genetic mapping; gene expression control;				99US-0134768.
KW	protein identification; signal transduction pathway;				99US-0134768.
KW	metabolic pathway; promoter; termination sequence; ss.				99US-0134768.
XX					99US-0134768.

```
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 20.5%; Score 381; DB 21; Length 1653;
Best Local Similarity 57.5%; Pred. No. 2.7e-79;
Matches 797; Conservative 0; Mismatches 560; Indels 30; Gaps 5;

Qy 171 TTCCCATATAGGAACCTTAACCTCTCGAACAGCAATCCACCGTTCTTCTTCCACGCA 230
Db 132 TACCGGTGATTGGTCACTCCGTCTTCTCAAAACACCGATTTCGCACATTCCTCTCC 191
Qy 231 TGTGAAACAGTACGGCAACG-----TGGTTTCCTCTGGTTCCGTTCACTCTGGCGG 284
Db 192 TCTCTCAATCCCTAAACAATGTCGATCTTCTCCCTCCGACTCGTAACCGACTCGTT 251
Qy 285 TTGTCATCTCTCTCCAAAGCATACCAAGATGTTTCCACCAACACGAGTTGCTTCGG 344
Db 252 TCGTGAACCTGTCACACTCGATCGCGGAGAAATGTTTCAACAAACCGACTCGTACTGG 311
Qy 345 CCAACCGGTACTCTCTCTCGGAAATATACATCTTCTACAAACACACCGTAGGCT 404
Db 312 CGAACAGACCAACTTCACTCTCGTAAACAGTTGCGTACGATTACACAACCATATCG 371
Qy 405 CCTGCTCCACGGGAGCACTGGCGCAACCTCCGGCGATACCGCCCTGGACGTCCTCT 464
Db 372 CAGCTTCTACGGTGACCACTGGCGTAACCTCCGGCGATCGCTCCGTCGAGATATTCT 431
Qy 465 CCACGAGCGGTCCACTCTCTCCGGAATCCGGAGCGACGACGAGACGCTCTGATCG 524
Db 432 CCAACCGGTCTCAATAGCTTTCTGTCTATTTCGTAAGACGAGATCCGACGACTGTGT 491
Qy 525 AGAGGTGTGTGCTGCGCAAGAACTCGAACGAGGAAGTTTCGCGGAGTGGAGATTAGTT 584
```

```
Db 492 TTGGTCTTTTACGTAACGTTT-----CACAAAGAGTTTGTGAAGTGGATATGAAT 542
Qy
585 CGATGTTTCAACGACTTAACCTTACAAACATAATAGGATCATATCGGGGAAGAGTTT 644
Db
543 CAATGTTATCTGACTTAACATTCACACACATTAAGAAATGGTGGCGGAACGTTACT 602
Qy
645 ACGGAGGAGAGTGAATGAAGAACGTTGAGGAAGCGGAGGAGTTTCAAGAGACTGTGA 704
Db
603 ACGGAGACGGTGTG-----AGGATGATCCGAGGCTAAACGTTGTCCGACGCTTAG 656
Qy
705 CAGAAATGTTGGAACACTCATGGCTTGCCTAACAGGAGAGATCACTTCCCTCAGT 764
Db
657 CGGATGGTGGCTGTGCTGGTCTGGAACGCTGTGATTAATACCGGTTTGGCGT 716
Qy
765 GGTTCGAATTTTCAAAATGCGAGAACGCTTTAAAGAGTATCAGTAAGAGGATGAATCCA 824
Db
717 TG-----GTTTCAGATTACGAACACGGGTAAAGAGTTAGCGGTAGGCTCGACGAGT 770
Qy
825 TCTTGAATAGATCTTTCATGAGAACCGTGCAGCAATGACCGCCAGAAATTCATGATCG 884
Db
771 TCTTGAAGGATTTGTTGATGAGAAACGAGACGCTAAGGAGAAAGGAAACACTATGATTG 830
Qy
885 ATCATCTCTCAACTCGAAGACCCAGCTCAGTACTACACTGACCAATCATCAAG 944
Db
831 ATCATCTGCTTACTCTCAAGATCAACACCGGATTAATCTCACTGATCGTATCATCAAG 890
Qy
945 GCCTTGTCTGGCCATGCTTTTGGTGGAACTGACTCATCACTGAGCTTTTAGAGTGGT 1004
Db
891 GAACATGCTTGTCTTGTACTAGCAGGACCCGACACATCAGCGGTTACGTTAGATGGG 950
Qy
1005 CATATCTAATTTAATGAATCACCAGAGGTTTGAAGAGGCAAGAGATGAATTTGACA 1064
Db
951 CATTGTGCAAGGTTTGAACCAATTCGGAGGTTATTGAACAAGGCGAGAGATGAATCGATA 1010
Qy
1065 CTCAGTGGGACAGACCGCTTGTAAATGAGTCAGACCTTCCAAACTTCCATATCTTA 1124
Db
1011 GAAGATCGGTTAGACAGGCTTAAGATGAATCAGATATCTCAACCTGCTTAATCTCC 1070
Qy
1125 GGAAGATCATCTTTGAGACACTTAGTGTGTACCCCGCGGCCCAATTTCTAATACCTCATG 1184
Db
1071 AAAACATTGTGCTGAAACGTTGCGCTTTATCTGCGGCTCCCATGCTTCTTCTCAAG 1130
Qy
1185 TGTCTTCAAGAGATTAATCAATTTGAAGATTAATCAATCCACGAGACACAAATTTGTATCA 1244
Db
1131 TTGCTCGAAGATTTGAAGTTGAGGATATGATATGCGCGTGGCAGCATACTATTGA 1190
Qy
1245 TTAATGTTGGGATCGACAGAGATCTCTAGTTGTGAATGATGCGCACATGCTTTAAAC 1304
Db
1191 CCAATGTGGGCTATACAGAGATCTCTAGCTATGGATGATCCAAATGAGCTTCAAGC 1250
Qy
1305 CTGAGAGTTTGTATGTGAAGGAGAGGAGAAAGTTGGTAGCAATTTGGCATGGGAAGAA 1364
Db
1251 CAGAGAGTTTGAAGAGGAGGAGAGCTCAGAGACTTATGCGGTTTGGGTTAGGAAGAA 1310
Qy
1365 GGGCTTCCAGGAGAACCCATGGCTATGCAAGTGTGAGCTTACCTTTTGGATTTGTA 1424
Db
1311 GGGGCTGTCTGGTCTTGGACTGCTCACCGGCTTATAAACCCTGACTCTTGGATCAATGA 1370
Qy
1425 TTCAATGTTTGTGCTGGAACAGAGTAAGTGAAGGAAAGCTTGTATGACAGAGAAACAATT 1484
Db
1371 TTCAGTGTTTGGAATGGA-----GAAGATTGAGAGAGAGTGGATATGAGTGAAGGAAAG 1427
Qy
1485 GGATCAGCTTGTCAAGGTTAATTCATTTGGAGGCGCATGTGCAAGGCTCGCCCACTGCCA 1544
Db
1428 GTGTTACAAATCCCTAAAGCCAGCCTTTAGAAGCCATGTGACAGAGCAGCTTCCCTCTGTG 1487
Qy
1545 CTAATAAT 1551
Db
1488 TTAATAAT 1494
```

RESULT 15
ABZ12256

```
ID ABZ12256 standard; DNA; 1503 BP.
XX
AC ABZ12256;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 61.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US26685.
XX
PR 24-AUG-2000; 2000US-227866P.
PR 26-JAN-2001; 2001US-26467P.
PR 22-JUN-2001; 2001US-30011P.
XX
(SCRI ) SCRIPPS RES INST.
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed
and producing plants with increased tolerance to these abiotic stresses
XX
PS Claim 144; SEQ ID NO 61; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
cell has been exposed, comprising:
CC (a) contacting nucleic acid representative of expressed polynucleotides
in the plant cell with an array or probes representative of the plant
cell genome; and
CC (b) detecting a profile of expressed polynucleotides in the plant cell
characteristic of a stress response. The method is useful in the
production of transgenic plants, cells and seeds and in producing plants
with increased tolerance to abiotic stress. The present sequence is that
of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
in methods of the invention.
CC Note: The sequence data for this patent is not represented in the printed
specification but is based on sequence information supplied to Derwent by
the European Patent Office.
XX
SQ Sequence 1503 BP; 397 A; 348 C; 374 G; 384 T; 0 other;
Query Match 20.3%; Score 377; DB 24; Length 1503;
Best Local Similarity 58.1%; Pred. No. 2.3e-78;
Matches 810; Conservative 0; Mismatches 550; Indels 33; Gaps 7;
Qy 171 TTCCCAATAATAGGAACCTTAACCTCTCGAAGCAGCAATCCACCGTTTCTTCCACGCA 230
Db 113 TACCGGTGATCGGTACCTCCGCTTCTCAAGCGCCACTCCACCGAGTTTCTATCCG 172
Qy 231 TGTGGAACAGTACGCAACG-----TGGTTTCCCTCTGGTTCGGTTACGCTGCGCG 284
Db 173 TCTCTGAGTCCCTTAGCGGATGCTCCGATTAATCTCCCTGCTCTCGGCAACCGACTCGTTT 232
Qy 285 TTGTCACTCTCCCTCCACAGCATACCAAGATGCTTCCACCAACAGCAGTTGCTTGG 344
Db 233 TCGTTGTCTCTTCACTCACTCACTCGCGGAGATGCTTCAAGAGAGAGCTGCTACTCG 292
Qy 345 CCAACCGGCTACCTTCTCTCTCGGGAATAATACATCTTCTACAAACACACCGTAGGCT 404
Db 293 CGAACCGCTTCACTCTCTCTTGTTCAAACACACATCTCATATGGCTGCACACGGTTGTA 352
Qy 405 CTGTCTCCACGCGGAGCACTGCGGCAACCTCGCGCGATCAACCGCCCTTGAGAGCTCTCT 464
```

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2003, 04:47:10 ; Search time 1657 Seconds
(without alignments)
3008.731 Million cell updates/sec

Title: US-09-868-546A-1

Perfect score: 1859
Sequence: 1 Gaaacactgacagacagca.....aataaattttttactttc 1859

Scoring table: IDENTITY NUC

Gapop 10.0 , Capext 1.0

Searched: 1792395 seqs, 1340900451 residues

Total number of hits satisfying chosen parameters: 3584790

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMP.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMP.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMP.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMP.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMP.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMP.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMP.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMP.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMP.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMP.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	393.8	21.2	1488	10	US-09-938-842A-1188
2	386.6	20.8	1494	10	US-09-938-842A-2278
3	383.8	20.6	1479	10	US-09-938-842A-2172
4	383.2	20.6	1503	10	US-09-938-842A-1473
5	377	20.3	1503	10	US-09-938-842A-61
6	361.8	19.5	1509	10	US-09-938-842A-835
7	350.6	18.9	1557	10	US-09-938-842A-1959
8	330.4	17.8	1548	10	US-09-938-842A-1541
9	202.8	10.9	722	9	US-09-770-149-144
10	192.6	10.4	923	12	US-10-259-165-565
11	189	10.2	975	12	US-10-259-165-433
12	189	10.2	978	12	US-10-259-165-101
13	182.2	9.8	447	15	US-10-174-693-221
14	181.6	9.8	1545	12	US-10-356-153-55
15	181.6	9.8	1545	14	US-10-142-231-55
16	173.4	9.3	981	12	US-10-259-165-743

c	17	168.6	9.1	475	9	US-09-770-444-62	Sequence 62, Appl
	18	159.4	8.6	1530	12	US-10-097-559-12	Sequence 12, Appl
c	19	153.6	8.3	557	12	US-10-259-165-490	Sequence 490, Appl
	20	143.4	7.7	494	15	US-10-174-693-222	Sequence 222, Appl
c	21	139.4	7.5	495	10	US-09-924-035A-358	Sequence 358, Appl
	22	139.4	7.5	495	11	US-09-770-961-559	Sequence 559, Appl
c	23	139	7.5	1557	10	US-09-938-842A-2669	Sequence 2669, Appl
	24	128.6	6.9	1824	10	US-09-814-786-1	Sequence 1, Appl
	25	127.2	6.8	622	15	US-10-174-693-17	Sequence 17, Appl
	26	126.2	6.8	1575	10	US-09-938-842A-1997	Sequence 1997, Appl
	27	125.8	6.8	1509	12	US-10-259-165-145	Sequence 145, Appl
	28	125.2	6.7	670	12	US-10-259-165-568	Sequence 568, Appl
	29	121.8	6.6	1527	9	US-09-931-267-1	Sequence 1, Appl
	30	121.8	6.6	1527	9	US-09-931-267-2	Sequence 2, Appl
	31	121.8	6.6	1549	9	US-09-931-267-3	Sequence 3, Appl
c	32	117	6.3	508	11	US-09-770-961-255	Sequence 255, Appl
	33	117	6.3	2174	10	US-09-814-786-63	Sequence 63, Appl
c	34	116.6	6.3	506	10	US-09-924-035A-188	Sequence 188, Appl
	35	114.6	6.2	1929	11	US-09-899-842-1	Sequence 1, Appl
	36	110.6	5.9	1812	12	US-10-171-174A-6	Sequence 6, Appl
	37	108.8	5.9	1708	9	US-09-796-256A-1	Sequence 1, Appl
	38	108.4	5.8	1539	12	US-10-356-153-53	Sequence 53, Appl
	39	108.4	5.8	1539	14	US-10-142-231-53	Sequence 53, Appl
	40	107.8	5.8	621	12	US-10-259-165-675	Sequence 675, Appl
c	41	105.8	5.7	738	9	US-09-770-149-70	Sequence 70, Appl
	42	102.4	5.5	806	9	US-09-770-445-795	Sequence 795, Appl
c	43	100.8	5.4	1539	10	US-09-938-842A-199	Sequence 199, Appl
	44	98.2	5.3	1491	10	US-09-938-842A-2690	Sequence 2690, Appl
	45	98	5.3	1933	13	US-10-067-534-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-938-842A-1188
; Sequence 1188, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harter, Jeff
; APPLICANT: Krups, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1188
; LENGTH: 1488
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1188

Query Match 21.2%; Score 393.8; DB 10; Length 1488;
Best Local Similarity 57.4%; Pred. No. 3.2e-95;
Matches 801; Conservative 0; Mismatches 567; Indels 27; Gaps 4;

QY	162	CTCCCTCTTCCCATATAGAACCTTAACCTCTCGAACAGCCATCCACCGTTTCT	221
DB	107	CTTACCTTTTACCGTTATCGCCACCTCCACCTCTCAAGAACCGTGCACCAACAT	166
QY	222	TCCAACGATGTGGAACAGTACGCAACG-----TGGTTCCCTCTGTTGGTTCAC	275
DB	167	TCCACTCCATCTTAAGTCCCTTGGAAATGTCCTCAATTTCCACTTCGCTGGAAACC	226

Qy 276 GTCTGGCGGTTGTCTCTCTCCAAACAGCATACCAAGAAATGTTTCCACAAACACGACG 335
Db |||||
Qy 227 GCCTCGTTTACGTCTCTCTCCACATAGCCGAAGAATGTTTCCACAAACACGACG 286
Db |||||
Qy 336 TTGSCCTTGGCCAAACGGGTACTTCTCTCGGGAATAATACATCTTCTAACAACACCA 395
Db |||||
Qy 287 TCGTTCCTTGGAAACGGTCTGATATCATCATGCGCAACACGTCGGCTATAACTTTACCA 346
Db |||||
Qy 396 CCGTAGGCTCTGCTCCACCGCGAGCACTGGCCCAACCTCCGCGCATCACCGCCCTCG 455
Db |||||
Qy 347 ATATGATTGACATCTTACCGCGACCACTGGAGGAATCTCCGCGCATGCGCGCGCTCG 406
Db |||||
Qy 456 ACGTCTCTCCACGAGCGGTCCATCTCTTCTCGGAATCCGAGCGAGAGCAAGC 515
Db |||||
Qy 407 AGATAATTCTCTCATAGAAATCAGTACCTTTCTTCTATCCGTAAGAGAGAGATCCGAC 466
Db |||||
Qy 516 GTCTGATCAGAGGTTGGTGTGCGCAAGAACTCGAAGGAGAGAGATTTGCGCGAGTGG 575
Db |||||
Qy 467 GGCTCATAAACACATCTCTTAGAGACT-----CCTTGACCGGATTTGTTGAAGTGG 517
Db |||||
Qy 576 AGATTAGTTCGATGTTCAACGACTTAACCTTACAAACAATAGGATGATATCGGGGA 635
Db |||||
Qy 518 AGTTGAATCGTTTATTAACCAACTTGGCATTCACAACATCATCATGATGGTAGCCGAA 577
Db |||||
Qy 636 AGAGGTTTACGAGAGAGAGTGAAGAAAGCTTGAAGAGCGAGGAGGAGTTTCAGAG 695
Db |||||
Qy 578 AACGATATTACGG-----TACCGGTACAGAAGCAATGATGAAGCCAAAGCTCGTAGGG 631
Db |||||
Qy 696 AGACTGTGACAGAAATGTTGGAATCATGCGCTTGGCTTAACAGGGAGATCACTTGCCT 755
Db |||||
Qy 632 AACTTATAGCGGAGATAATGCCGCGCGGTCTTGGAAATTTGGCTGATATCTTCCGT 691
Db |||||
Qy 756 TCCTCAGGTGTTCCGATTTTCAGAAATGTGGAAGAGCGCTTAAAGAGTATCAGTAAGAGGT 815
Db |||||
Qy 692 CCAATATATGGGTC-----ACAACTTTGAGAACCAAGACCAATCTTGGGGAATCGAC 745
Db |||||
Qy 816 ACGATTCCATCTGAATGAAGTCTCTCATGAGAACCGTGCAGCAATGACCGCCAGAAAT 875
Db |||||
Qy 746 TCGATAGATGCTCTGCAAAACTGTTGATGAGAAACGTCGAGAGAAAGAAAGGTCAAA 805
Db |||||
Qy 876 CCATGATGATCATCTCTCAAACTGCAAGAGACCCAGCTCAGTACTACATCAACCAAA 935
Db |||||
Qy 806 CTTTGATCGATCATTTGCTTCTTCCAAAGAACCGAACCCGAGTACTATACTGATGCA 865
Db |||||
Qy 936 TCATCAAGAGCGCTGCTCTGGCCATGCTTTTGTGGAATCTGACTCATCACTGGGACT 995
Db |||||
Qy 866 TCATCAAGAGATCATATGCTGCTTGGTCTTTCGGGGGAGAGATACATCTTCAGTACGT 925
Db |||||
Qy 996 TAGAGTGTCTAATCTAATTTAATGAATCAACAGAGGTGTTGAAGAGGCAAGAGATG 1055
Db |||||
Qy 926 TGGAAATGGCAATGTCAAAATTTGTGAACCATCCAGAGATACTTGAGAAAGCGAGAGCG 985
Db |||||
Qy 1056 AATTGGACACTCAAGTGGGCAAGACCGCTTGTAAATGAGTCAGACCTTCCAAACCTC 1115
Db |||||
Qy 986 AGATCGATGATAAATCGGTTGAGACCGGTTAGTTGAAGAAATCAGATATTTAATCTCC 1045
Db |||||
Qy 1116 CATATCTTAGAAGATCATCTTGAGACACTTAGGTTGTACCCCGCCGCCCAATTTCTAA 1175
Db |||||
Qy 1046 ATTATCTCCAAACATTTGTGTGAGAAACATTAAGTTTGTATCTCGGTTCCACTACTAC 1105
Db |||||
Qy 1176 TACCTCATGTCTTTCAGAGATATTAATTAAGAGGATTCATATATCCAGAGACACAA 1235
Db |||||
Qy 1106 TCCCTCATTTCTCATCGGATGAATGTAAGTGGCGGCTAGCATATGCGAGTCCGACGT 1165
Db |||||
Qy 1236 TTGTGATCATTAATGTTGGGCGATGAGAGATCTCTCAGTTGTGGAATATGCCACAT 1295
Db |||||
Qy 1166 TGTATTAAACACGTATGGGCGATGCTAGAGATCCAGGTTTATGGAAGAGCCAGAGA 1225
Db |||||
Qy 1296 GCTTTAAACCTTGAGAGGTTTGTATGGAAGAGAGAGAGAAAGTTTGTGTAGCATTTGGCA 1355
Db |||||
Qy 1226 GGTTCAGCCGAGAGGTTTCGAGAAAGAGAGAGGCTCTGAAGACTAATGCGGTTGGGA 1285
Db |||||
Qy 1356 TGGGAAGAGGGCTTGGCCAGAGAAACCATGGCTATGCAAAAGTGTACAGTTTACTTTGG 1415
Db |||||

Db 1286 TGGAGCGAGAGCTTGTCTCGAGCTGAGCTTTGGGAAGCGGTTAGTGAGCCTTGCCTTGT 1345
Qy |||||
Qy 1416 GATTGTTGATTCAATGTTTGTAGCTGGAACGAGTAAAGTGAAGAAAGCTTTGATATGACAG 1475
Db |||||
Qy 1346 GGTGTTGATTGATCTTTTCGAGTGGGAGAGAGTTGTCAGAACTTGTGGACATGACTG 1405
Db |||||
Qy 1476 AGAACAAATTTGATACCTTTGTCAAGGTTAAATTTCCATTGAGGCGCATGTGCAAGGCTCGCC 1535
Db |||||
Qy 1406 AAGGCGAAGGATCACTATGCTAAAGCTACTCCGTTGGAGCTATGTGCAAGGCAAGTG 1465
Qy |||||
Qy 1536 CACTTGCCTACTAAAA 1550
Db |||||
Qy 1466 CCAATTGTTGGTAAAA 1480
Db |||||

RESULT 2

US-09-938-842A-2278
; Sequence 2278, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2278
; LENGTH: 1494
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2278

Query Match 20.8%; Score 386.6; DB 10; Length 1494;
Best Local Similarity 57.0%; Pred. No. 2.8e-93;
Matches 839; Conservative 0; Mismatches 599; Indels 33; Gaps 6;

Qy 90 TTTTCTCTGTTCTTATCTCTCGGCGTGAAGTTTGTGTTTCCAAAGCAGAAATTCAGAAACA 149
Db 29 TTCTCTCTGTTGTTCTCTCTCAATTTACTTAATTGGAATACTCAAGCGAAAGCCAAATC 88
Qy 150 TACCACACAGTCTCTCC---TCCTCTTCCCATATAGGAACCTTAACTCTCGAAGCAGC 206
Db 89 TACTCTCGAGTCGGGATGTCATTACCGGTGATTGTCATCTCCGCTTCTCAAAACCAC 148
Qy 207 CAATCCACCGTTTCTTCCAAACGATGTCGAAACAGTACGGCAACG-----TGGTTTCCC 260
Db 149 CGATTTCATGCAATCTCTCTCTCAATCCCTAAACATGTCTCCGATCTTCTCTCCC 208
Qy 261 TCTGTTTCGGTTTCACTCTGCGGCTGTCATCTCTCTCCATCCCAAGAGCATACCAAGATGCT 320
Db 209 TCCGACTTCGGTACCAGCTGTTTTCGTGAACCTCGTCACACTCGATCGCGCGAAGATGTT 268
Qy 321 TCACCAAAACACGAGTTCCTTGGCCAAACCGGCTACTCTCTCTCTCGGGAATAATCATCT 380
Db 269 TCACCAAAACGAGCTGCTACTGCGGACAGACCAAACTTCATCTCTCGCTTAACACGTTG 328
Qy 381 TCTTAAACAAACCAACCGTAGGCTCTCTGCTCCCAACGCGAGCACTGCGCGCAACCTCGGCC 440
Db 329 CGTACGATTACAAACCATGATCGCAGCTTCTCTACGCTGACCACTGCGGTAACCTCGGCC 388
Qy 441 GCATCACCGGCTGGAGGCTCTCTCCACGAGCGGCTCCACTCTCTCTCGGGAATCCGGA 500
Db |||||

389 GCATCGGCTCGGTCGAGATATCTCCAAATCACCGTCTCAATAGCTTTTGTCTATTCTGTA 448
Db
501 GCGACGAGCAAGCGCTGTGATCGACAGGTGTGGTGGCCCAAGAACTCGAACGAGGAAG 560
Qy
449 AAGACGAGATCCGACGACTGTGTTCGTCTTCTCGGAAC-----TTTTCACAAG 499
Db
561 AGTTTCGGCGAGGAGATAGTTCGATGTTCAACGACTTAACCTTAACCAACATATGCA 620
Qy
500 AGTTTGTGAAGTGAATGAATCAATGTTATCTGACTTAACTTAACCAACATTTTAA 559
Db
621 GGATGATATCGGGGAAGAGGTTTACGGAGAGAGAGTGAAGATGAAGAAAGCTTGAAGAG 680
Qy
560 GAATGCTGGCCGAAACGTTACTACGGAGACGTTGTG-----AGGATGATCCGGAGG 613
Db
681 CGAGGAGTTTCAGAGAGACTGTGACAGAAATGTTGGAACCTCATGGCTTGGCTTAACAAGG 740
Qy
614 CTAACAGTGTCCGGCAGCTTATAGCGGATGTGGTGTGTGTGCTGGAACGCTG 673
Db
741 GAGATCACTTGCCTTTCCTCAGGTGTTGATTTTCAGAAATGTCGAGAGCGCTTAAAGA 800
Qy
674 TTGATTACTTACCGGTTTTCGGTTGG-----TTTCAGATTACGAGACCGGTTAAGA 727
Db
801 GTATCAGTAAGAGGTACGATTCATCTTGAATAAGATCCTTCATGAGAACCGTGCAGCA 860
Qy
728 AGTTGGCGGTAGCTCGACGAGTTCTTGCAGGATTTGGTGTGATGAGAAACGAGAAGCTA 787
Db
861 ATGACGCCAGAAATCCATGATCGATCATCTCTCAAACTGCAAGAGACCCAGCTTCAGT 920
Qy
788 AGGAGAAAGGAACACTATGATCGATCATCTTACTCTGCAAGAATCCCAACCGGATT 847
Db
921 ACTACACTACCAAAATCATCAAGCGCTTGTCTGTGGCCATGCTTTTGTGGTGAACGTACT 980
Qy
848 ACTTACCGATCGTATCATTAAGAGAAACATGCTCGCTTTGTATACGAGGAGCCGACA 907
Db
981 CATCAACTGGGACTTTAGAGTGTCTATCTAATTTATTTGAATCAACCCAGAGGTGTCA 1040
Qy
908 CATCAGCGGTAGCTTAGAATGGGCATTTGCGACGTTGTAACCATCCGATGTATTGA 967
Db
1041 AGAAGCAAGAGATGAATTTGACACTCAAGTGGGCAAGACCGCTTTGTTAAATGATGTCAG 1100
Qy
968 ACAAGCGAGAGATGAATCGATAGAAAGATAGTTTACAGAGCTTATGGATGAATCAG 1027
Db
1101 ACCTTCCAAATCTCATATCTTAGNAGATCATCTTGGACACTTAGTTGTACCCCC 1160
Qy
1028 ATATCTCAACCTGCTTATCTCCAAACATTTGTCTGAAACGTTGCGCTTTATCCTG 1087
Db
1161 CGCCCCCAATTTCAATACCTCATGTCTTTCAGAGATATTACAATTTGAAGGATTTCAATA 1220
Qy
1088 CGGCTCCATGCTTCTTCTCAGCTTGCCTCGGAGATTTGTAAGTTGAGGATACGATA 1147
Db
1221 TCCCGAGACACAATTTGATCATTAATGGTTGGGGCATGCGAGAGATTCCTCAGTTGT 1280
Qy
1148 TGCCGGTGGCAGCACTATTACCAATATTGTGGGCTATACACAGAGATCTCTCAGCTAT 1207
Db
1281 GGAATCATCCCATCTTTAACTTGAAGGTTTGTATGTAAGAGAGAGAGAGAAAGT 1340
Qy
1208 GGGATGATCCATGAGCTTCAAGCCAGAGAGGTTTGAGAAAGAGAGAGAGCTCAGAGC 1267
Db
1341 TGGTAGCATTTGGCATGGGAAGAGGCTTGGCCAGGAGAACCCATGGCTTATGCCAAGTG 1400
Qy
1268 TAATGCGGTTTGGTTAGGAGAGAGGCGGTGCTCGTCTGCACTGGCTCATCGGCTTA 1327
Db
1401 TCAGCTTTACTTTGGGATTTGATTAATCAATGTTTTCAGTGGAAACGAGTAAGTGAAGAA 1460
Qy
1328 TAAACCTGACTCTTGGATCATTTGATTTGAGTTTGGAAATGGGA---GAAGATTTGAGAGA 1384
Db
1461 AGCTTGTATGACAGAGAACAAATTTGATCACCTTGTCAAGTTTAATTCATTGGAGGCCA 1520
Qy
1385 AGTGTGATAGTGAAGGCAAGGTTTACATGCTTAAAGCCAAAGCCCTTTTGAAGCCA 1444
Db
1521 TGTGCAAGGCTCGCCCACTTGGCACTTAAAT 1551
Qy
1445 TGTGCAAGACCTCTGTTGTTAAAT 1475
Db

RESULT 3

US-09-938-842A-2172
; Sequence 2172, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2172
; LENGTH: 1479
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2172

Query Match 20.6%; Score 383.8; DB 10; Length 1479;

Best Local Similarity 57.2%; Pred. No. 1.6e-92; Mismatches 557; Indels 36; Gaps 4;

Matches 794; Conservative 0;

Qy 171 TTCCCATATAGGAAACCTTAACCTCTCGAACACGCAATCCACCGTTTCTTCCCAACGCA 230
Db
116 TGCCGGTGTGCGACACCTCCATCTCTCAAGCCACCGCTTCAACGTACCTTCCCTTCC 175
Qy 231 TGTGAAACAGTACGGCAA-----CGTGGTTTCCCTCTGGTTCCGTTACGCTCTGGCGG 284
Db 176 TTTCAAATCCATCGAAATGCTCGGTCTTCCATCTCCGACTCGGAAATCGCCTCGTTT 235
Qy 285 TTGTCTATCTCTTCCAAACAGATACCAAGATGTTTCCAAACACGAGAGAGAGAGAGAG 344
Db 236 ATGTCATCTCTTACGTTTCCATCGCTGAAGATGTTTCAAAAGAACAGATGCTGTTCTCG 295
Qy 345 CCAACCGGCTACCTTCTCTCGGAAATATACATCTTACAAACACACACCGTAGGCT 404
Db 296 CGAACCGCCCAAGTTCAACATAAGTAAAGCACTCGGCTACAAACGCCACCTTACTTCT 355
Qy 405 CTTGCTCCCAAGCGGAGCACTGGCGCAACCTCCCGCGCATCACCGCCCTGGAGAGCTCTCT 464
Db 356 CGGCATCTTACGGCGCATTTGGAGGAACCTCCCGCGCATAGCCCGCTCGAGATATTTCT 415
Qy 465 CCACGACGCGGTCCACCTCTTCTCGGAATCCCGAGGACGAGAGAGAGAGAGAGAGAG 524
Db 416 CCATCTATAGACTCAATTTCTGTTTCTGTATATTCGTAAGACGAGATCCGACGCTCATTT 475
Qy 525 AGAGTTTGTGTGTGTCGCAAGAACTCGAACGAGAGAGAGTTTTCGCGAGTGGAGATTAGTT 584
Db 476 CACATCTCTCTGTTGATT-----CCTTACACGGATTGTGAGAGTAGAGATGAAAA 526
Qy 585 CGATGTTCAACGACTTAATCTTAAACAACTAATGAGGATGATATCGGGGAGAGAGTTT 644
Db 527 CATTTGTTAAACAAATTTGGCATCCCAACCAACCATCAGAAATGTTGGCCGCGAGAGATA 586
Qy 645 ACGGAGAGAGAGTGTGATGAAGAGCTTTGAGGAGCGGAGAGTTTCAGAGAGCTGTGA 704
Db 587 TCGGTGAGG-----ACAACGATGACGCTAACTCGTGAAGAACCTTTGTGT 631
Qy 705 CAGAAATGTTGGAACTCATATGGGCTTGGCTTAAACAGGAGATCACCTTGGCTTTCTCAGGT 764
Db 632 CGGAGCGGTGACACGCGCGGTGAGGAAACCCCATTTGATTATCTTTCCATTTTACGTT 691

QY 765 GGTTCGATTTTCAGATGTTGGAGCGCTTAAGAGATATCAGTAAGAGGTACGATTCGA 824
Db |||||
QY 692 GGGTCTC-----GAGTTATGAAAACGAATCAAGAAATTTGGGAAATAGGTTTGAATCGT 745
Db |||||
QY 825 TCTTGAATAAGATCCTTCATGAGAACCGTGCAGCAATGACCGCAGATTTCCATGATCG 884
Db |||||
QY 746 TTTTGCAGAAATTAGTCACGAAACCTGGGAGAGGAAAAGGTGAACATGATCG 805
Db |||||
QY 885 ATCATCTCTCAAACTGCAAGAGAGCCAGCTCTAGTACTACATGACCAAAATCATCAAG 944
Db |||||
QY 806 ATCACTTCTGCTCTCCAGAGACATTCACCTGATTAATACGATGTCATCATCAAG 865
Db |||||
QY 945 GCCITGCTCTGCCATGCTTTTGGTGAATCTACTCATCACTGAGGACTTTAGAGTGT 1004
Db |||||
QY 866 GAATCATACTTACCTGATAATTGCGGRCAGATACGTCATCAGTAACACTAGAAATGGG 925
Db |||||
QY 1005 CATATCTAAATTATTAATCAACCAGAGAGTGTGAAGAGCGCAAGAGATGAATGGACA 1064
Db |||||
QY 926 CAATGTCAAATCTGTTGAACCATCCAGAAATACTTTAGAAAGCGAGATGGAATCGATG 985
Db |||||
QY 1065 CTCAGTGGGCAAGACCGCTTGTAAATGATGACAGCTTCCAAAATTTCCATATCTTA 1124
Db |||||
QY 986 AAAAAGTCGGTTAGACCGAATTAGACGAATCGACATTTGTAATCTCTCTTATCTCC 1045
Db |||||
QY 1125 GGAAGATCATCTTTGAGACACTTAGGTTGTACCCCGGCCCAATTTCTAATACCTCATG 1184
Db |||||
QY 1046 AAGCAITGTATFTGGAACACCTACGATGTATACCGGCGAGTCCCACTACTACTCTCAAT 1105
Db |||||
QY 1185 TGTCTTCAGAAATATTACAATTGAAGGATTCATATATCCACGAGACACAAATTTGTATCA 1244
Db |||||
QY 1106 TGTCAATCAGAAATGTTAAAGTTGGAGCTAGGATATACCAAGTTGGAACATGTTATGA 1165
Db |||||
QY 1245 TTAATGTTGGGCGATGAGAGAGATCTCTAGTTGTGGAATGATGCCACATGCTTTAAAC 1304
Db |||||
QY 1166 CCAACGCTATGAGATGATGAGATCCAGAGGATGGAAGATCTCTGAGATATTCAAAAC 1225
Db |||||
QY 1305 CTGAGAGTTTGTGTGAAGAGAGAGAAAGTTGGTAGCATTTGGCATGGAGAA 1364
Db |||||
QY 1226 CAGAAAGATTTGAAAAGAGAGAGGCTGAGAGCTAATCTCATTTGGGATGGAGAA 1285
Db |||||
QY 1365 GGGCTTGCCAGGAGAACCCATGGCTATGCAAAGTGTGAGCTTTACTTTGGGATTTTGA 1424
Db |||||
QY 1286 GAGCTTGTCTGGAGCGGCTAGCTCATCGCTTAATAAACAGGCTCTTGGAGTTTGG 1345
Db |||||
QY 1425 TTCAATGTTTGTGCTGGAACAGAGTAAGTAGGAGAAAGCTTGTATGACAGAGAACAT 1484
Db |||||
QY 1346 TTCAATGTTTGTGTTGGAAGAGTGTGAGGATTTTGTGACATGACCGAAGACAAAG 1405
Db |||||
QY 1485 GGATCACTTGTCAAGGTTAATTCATTTGGAGGCGCATGTCAGAGCTCGCCCATTTGCCA 1544
Db |||||
QY 1406 GAGCCCATGTCCTCAAGCTATACCATTAAGAGCCATGTGCAAGACGTTCTTATTTG 1465
Db |||||
QY 1545 CTAAAT 1551
Db |||||
QY 1466 ATAAAT 1472
Db |||||

RESULT 4

US-09-938-842A-1473
; Sequence 1473, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1473
; LENGTH: 1503
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1473

Query Match 20.6%; Score 383.2; DB 10; Length 1503;
Best Local Similarity 56.5%; Pred. No. 2.3e-92;
Matches 847; Conservative 0; Mismatches 618; Indels 33; Gaps 6;

QY 85 TGTCTCTTTCTCTGTTTCTATTCCTCGGCGGTGAAGTTTGTTCCTCAAGCAGAAAAATTGAG 144
Db 18 TCTCTCTCTCTCTATTTCTTGAATTAAGTTTACAAATTTCTCTACTCCAAAACGCAAGTTT 77
QY 145 AAACATACACACAGGTCCTCTCTCTCCATATATAGAAACCTTAACCTCTCGAACA 204
Db 78 CAACTTCTCTCAGGACCAACCGTCAGACCTTCGTCGGCCATCTCCACCTCATGAAC 137
QY 205 GCCAATCCACCGTTCTTCCACGCACTGTCGAAACAGTAGCGCAACGTTGTTCCCTCTG 264
Db 138 GCCGATCCACCGTCTCTCCACGCTACTCCACCAATAGCGCCCAATCTTTTCCCTCCG 197
QY 265 GTTCGGTTCACTGCTGCGGCTGTCTCTCTCCACAGCATATCAAGAATGCTTTCAC 324
Db 198 TTTTCGGTCCGAGCGGTCTGCTGCTATTAATATCCCTTCTCTCGCCCAAGATCTTTCAC 257
QY 325 ---CAAAACAGAGTTGCTTGGCCAAACGGGTACTCTCTCTCGGGAATAATACATCTT 381
Db 258 CGGCCAAAACGATCTCTCTCTCCAGCGCGCGCTCCAACTCACCGCCAAATACGTCG 317
QY 382 CTACAAACACACACCGTAGGCTCTCTCCACGCGGAGCACTGGCGCAACCTCCGCGG 441
Db 318 ATACACACACACCGTTGTTGTTACAGCTCTTACGGGGACCACTGCGGTAACTCTCGCG 377
QY 442 CATCACCGCTCGAGCTCTCTCTCCACGAGCGCTCTCTCTCTCGGGAATTCGGAG 501
Db 378 CATTTGCTCCAAAGATTTCTCTCTCTCACGCTCTCATCACTTCCACACATCCGCA 437
QY 502 CGAGGAGACGAGCGTCTGATGAGAGGTTGTTGTCGGCCAGAACTCGAACAGGAA-- 559
Db 438 AGACGAGATTTACGAGTCTCTCACGCTCTCACGCTTACACACAACTCCCAACGAAAG 497
QY 560 ---GAGTTTGGCGAGTGGAGATTAGTTCGATGTTCAACGACTTAACCTTACAAACAT 615
Db 498 CAACGATTTCACTCATCGAGTAGAACGCTCTTATCTGATCTAACATTCACACAT 557
QY 616 AATGAGATGATATCGGGGAAGAGGTTTACGAGAGGAGAGTGAGATGAAGAACGTTGA 675
Db 558 AGTAAGAAATGGTTACAGGGAAGAGATATTACGCGCAGCAGC-----TCAACAAACAA 611
QY 676 GGAAGCGAGGAGTTTACAGAGACTGTGACAGAAATGTTGGAATCATGCGGCTTGGTAA 735
Db 612 AGAAGCAGAGCTATTCAAGAAAGCTAGTCTACGATATGCGCATATAGCGCGCTAATCA 671
QY 736 CAAGGGAGATCACTTGCTTTTCTCAGGTGTTTTCGATTTTTCAGAAATGTGAGAGCGCTT 795
Db 672 TTCGCTGATTTACTTACCAATATAAATTTTCGGAA---ACAAATTCAGAAAGAGT 728
QY 796 AAAGAGTATCAGTAAGAGGTAACGATTCATTTTGAATTAAGATCTTCTCATGAGAACCGTGC 855
Db 729 TAAAGCTATAGGCAATCAATGATGATATTTTGCAGCGTTTGTAGATGAAT---GTAG 785
QY 856 CAGCAATGACCGCCAGAAATTCATGATCGATCATCTCTCAAACTGCAAGAGACCGAC 915
Db 786 AAGAGATAAAGAGAGGTAACAAATGTTGTCATTTGATCTCTCTACAAACAAACAAAC 845
QY 916 TCAGTACTACACTGACCAATCATCAAGGCTTGTCTGCGCCATGCTTTTGTGTGAAC 975
Db |||||

QY	1242	TCATTAAATGGTTGGGGCATGAGAGAGATCCTCAGTTGTGGAATGATGCCACATGCCTTTA	1301
Db	1172	TGACTAATGATGCGGCAATACATAGGATCCTAAGATATGGATGATCCAACGAGCTTCA	1231
QY	1302	AACCTGAGAGGTTTGATGTGGAAGGAGAGAGAAAAGTTTGGTAGCATTTGGCATGGGAA	1361
Db	1232	AACAGAGAGGTTTCGAGAAAAGAGGGAGGCTCAGAAAGCTATTGGGGTTTTGGGCTAGGAA	1291
QY	1362	GAAGGGCTTCGCCAGGAGAAACCCATGGCTATGCAAAAGTGTGAGCTTTTACTTTTGGGATTGT	1421
Db	1292	GAAGGGCGTGTCTGGGTCCGGTCTGCTCAACGGCTAGCGAGTTTGACTATCGGGTCTT	1351
QY	1422	TGATTTCAATGTTTTGATCGTGAACACGAGTAAGTAGAGGAAAAGCTTTGATATGAC---	1478
Db	1352	TGATTTCAATGTTTTGATGGGAGAGGATAGGAGAAAGAGGTGATGATGACTGAAGGTG	1411
QY	1479	ACAATTGGATCACCTTCTCAAGGTTAATTCATTGGAGGCCATGTGCNAGGCTGCCCCAC	1538
Db	1412	GAGGAGGAGTCATAATGCCAAAGCTATACCGTTGGTAGCCATGTGAAAAGCACGCCCG	1471
QY	1539	TTGCCACTAAAAAT 1551	
Db	1472	TTGTTGTGTAAGAT 1484	

RESULT 6

US-09-938-842A-835

; Sequence 835, Application US/09938842A

; Patent No. US20020160378A1

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; TITLE OF INVENTION: SAME, AND METHODS OF USE

; FILE REFERENCE: SCRIP1300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 835

; LENGTH: 1509

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-835

Query Match	19.5%	Score 361.8	DB 10	Length 1509
Best Local Similarity	57.6%	Prod. No. 1.4e-86		
Matches 755	Conservative 0%	Mismatches 532	Indels 24	Gaps 5
Qy	245	GGCAAGTGGTTTCCTCTGTTCCGTTACAGTCTGGCGGTGTGTCATCTCTCCAA--	302	
Db	211	GGCGGGTGATGTCCTCCGACTGGATCCGCTTAGTCTACGTGGTGTGTGCATTAAG	270	
Qy	303	-CAGCATACCAAGTAATGTTTCACCAAAACGAGCGTTGCTTGGCCAAACGGGTACTTCT	361	
Db	271	GTTTCGGCAGAGGAGTGCTTTGGCAAGAACGACGTCGTTCAGTTAACCGGCCACAGTG	330	
Qy	362	CTCTCGGGAATAATACATCTTCTACAAACAAACACCGTAGGTCCTGTCTCCACGGCCAG	421	
Db	331	ATCATCGGAACAACATGTGGTTTATACAATAACAACATGATCGCTGCACCTTACGGTGAT	390	
Qy	422	CAC'TGGCGCAACTTCGGCCGATACACGCCCTTGGAGTCCTCTCCACGAGGGGTCCAC	481	
Db	391	CAC'TGGCGTAACTTCGGTCGGCTCTGCACTATCGAGATCTTTTCAACAACACGGCTTAAC	450	
Qy	482	TCCTTCTCCGGAATCCCGAGCGACGAGCAAGAGGTCTGTATGCAGAGGTGGTGTGGCC	541	

RESULT 7

[illegible]

US-09-938-842A-1959
 ; Sequence 1959, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIP1300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1959
 ; LENGTH: 1557
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-1959

Qy	743	GATC	ACTTGCTTTCTC	CAGGTGGTTTCG	ATTTTC	GAAATGTG	GAGAACGC	CTTAA	AGAGT	802		
Db	679	GA	TTTTTTCGCGCTCT	CAGGTGGTTTCG	ACTA	CAAAAGGTTT	TGCTCAAG	GAGCAAA	GAGG	738		
Qy	803	ATCAGT	AAAGGTTACG	ATCTCCATCT	TGAA	TAAGATCCTT	CATCAGAACCG	TGC	CCAGCAAT	862		
Db	739	ATTGG	AAAGGATGG	ATAGTTTATTA	CAAGGGTTTT	TGGATGAG	CAATAGAGCT	TAA	CNA	798		
Qy	863	GACCG	-----	CCAGAA	TTCCATGATCG	ATCATCTCT	CAAACTG	CAAGAGACCC	CAG	913		
Db	799	GACAG	ATTGGAGTTT	AAAGAAC	CCATGATG	CTCATTTACT	TGATTTCT	CAAGAAAA	AAGAG	858		
Qy	914	CCTCAG	TACTACTG	ACCAAAATCAT	CAAAAGCC	TGCTTG	CGCCATGCTTT	TGGTGA	973			
Db	859	CCTC	ATAACTACAGT	GATCAAA	CCATCAAAAGG	CTTTATCT	GTATGATGGTAG	GAGG	918			
Qy	974	ACTG	ACTCACTCA	CTGGACTTTT	PAGAGTGGT	CATTAATCT	TAATTTATTTGA	TATCA	CCACGAG	1033		
Db	919	ACAG	ATACATCG	GGCTTAACCG	TGGAA	TGGCAATGT	CAAAATTTATTTGA	TATCAT	CCACAA	978		
Qy	1034	GTGTT	GAAAGGCA	AGAGATGA	ATTGG	ACACTCAAG	TGGGCA	CA	-----	GACCGC	1084	
Db	979	ATAC	TAGAGACCA	CGAGACAAA	CAITGT	ACTCAAA	TGGAAAA	CATCA	AGTAGT	CGTCGT	1038	
Qy	1085	TTGTT	TAANTCAGT	CAGACCTT	CAAAA	ACTTCC	ATATCTTAG	GAGAGATCA	CTCTT	GAGACA	1144	
Db	1039	TTGTT	GAAAGAA	GAAGATTT	TGTPAA	CA	TGAACTGT	GAAAAATGTT	GTG	CGGAGACA	1098	
Qy	1145	CTTAG	GTGTAC	CCCCCGGCCAA	TTCTA	ATACCTCAT	GTGCTTC	CAGAAG	ATATTA	TACA	1204	
Db	1099	CTA	AGACTTTAC	CCGTGG	CAACACTT	ATGGTTCC	TATGTC	CACTTCTG	ACTGTG	TA	1158	
Qy	1205	ATTGA	AGATTCA	ATPCCA	CGAGACACA	AAATTTG	TGATCATTAAT	TGTTGGG	CA	TGCAG	1264	
Db	1159	ATCG	TGCAATCA	ACGTTCC	ACGTGAT	CAATCGT	GTGGTTAACT	TGTGGG	CTAT	ACAT	1218	
Qy	1265	AGAG	ACTCAG	TGTGG	ATGATGCC	ATGCTTTAA	ACCTCAG	AGGTTTG	TATG	-----	1319	
Db	1219	CGAG	ATCAAG	CGTGTGG	ATGATCC	AACTCAIT	TAAGCCG	GAGAGATTT	GA	GGAGT	1278	
Qy	1320	-----	TGGA	AGGAGAG	GAAAAAG	TGGTAG	CAATTTGCG	ATTTGCG	CAAGAA	GGGCTTGC	1372	
Db	1279	GATC	AGTTTGG	ACATTA	CAATGTTAA	ATGATCG	CGGTTT	TGSGTTAG	CAAG	CGGCTTGT	1338	
Qy	1373	CGAG	GAGACCC	ATGGCTAT	GAAAGT	GTCA	GCTTTACTTT	TGGGATTT	GATTCA	TATGT	1432	
Db	1339	CCTG	GGTTGATTT	TGGCTAA	CCGGGTTGTT	TGGGTTATTA	TATGGG	TTCCG	ATTA	CTAGT	1398	
Qy	1433	TTTG	ACTTGG	AAAA	CGAGTAA	GTAGG	AAAAAGCTT	GATATG	ACAG	AGAA	CAATTTGG	1492
Db	1399	TTTG	AGTTGG	AAAG	ACGGCT	CGAG	AGGTCA	GGTTGATAT	GTACT	GA	AGTCCAG	1458
Qy	1493	TTGT	CAAGGTTAA	TTCC	ATTGG	AGGCGA	TGTG	CAAGG	CTCG	1533		
Db	1459	TTGCT	TAAGCTGA	ACCAATTT	TGGTTGT	CACG	CTG	CAGAA	CACG	1499		

RESULT 8
US-09-938-842A-1541
; Sequence 1541, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1541
; LENGTH: 1548
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1541

Query Match 17.8%; Score 330.4; DB 10; Length 1548;
Best Local Similarity 54.5%; Pred. No. 4.1e-78;
Matches 805; Conservative 0; Mismatches 641; Indels 30; Gaps 6;

QY 86 GTCCTTTCTGTTCTATTCTCTCGGCGTGAAGTTGTTTTCAGAGCAAAATTTGAGA 145
Db GTCTTTCTGTTCTATTCTCTCGGCGTGAAGTTGTTTTCAGAGCAAAATTTGAGA 145
QY 76 GCCTTCTCTCTTTCTATTCTGTTAAACATTTTGTGTAACGGCTTAGCAAGTTTAAAC 135
Db GCTTCTCTCTTTCTATTCTGTTAAACATTTTGTGTAACGGCTTAGCAAGTTTAAAC 135
QY 146 AACATACCAACAGTCTCTCTCTCTCTCCATATAGGAAACCTTAACCTCTCTCGAACAG 205
Db AACATACCAACAGTCTCTCTCTCTCTCCATATAGGAAACCTTAACCTCTCTCGAACAG 205
QY 136 CTACCAACCAAGTCGGGTAGCCCTGCTGTTGATGGACACCTTCACCTCTCTGAAACTC 195
Db CTACCAACCAAGTCGGGTAGCCCTGCTGTTGATGGACACCTTCACCTCTCTGAAACTC 195
QY 206 CCAATCCACCGTTTCTTCCAAACGATGTCGAAACAGTACGGCAACG-----TGGTTTCC 259
Db CCAATCCACCGTTTCTTCCAAACGATGTCGAAACAGTACGGCAACG-----TGGTTTCC 259
QY 260 CTCGTGTTGGTTCAAGTCTGCGGTTGTCATCTCTCTCCACAGATACCAAGATGC 319
Db CTCGTGTTGGTTCAAGTCTGCGGTTGTCATCTCTCTCCACAGATACCAAGATGC 319
QY 256 CTCGGCTCGGAAACCGTCTCACGGTTGTAGTCTCTTCTATACCTCCATCGCTGAAGAATGT 315
Db CTCGGCTCGGAAACCGTCTCACGGTTGTAGTCTCTTCTATACCTCCATCGCTGAAGAATGT 315
QY 320 TTACCAAAACACAGAGTGTCTTGGCCAAACCGGTACCTCTCTCTCGGAAATATCATC 379
Db TTACCAAAACACAGAGTGTCTTGGCCAAACCGGTACCTCTCTCTCGGAAATATCATC 379
QY 316 TTTACAAAACAGCATGTTTTCGCAATAGCCAGAGTTAATCTCGGGAACATATT 375
Db TTTACAAAACAGCATGTTTTCGCAATAGCCAGAGTTAATCTCGGGAACATATT 375
QY 380 TTTACAAACACACCGTAGGTCTCTGTCCTCCACGGCAGCATCTGGCGCAACCTCCGC 439
Db TTTACAAACACACCGTAGGTCTCTGTCCTCCACGGCAGCATCTGGCGCAACCTCCGC 439
QY 376 GAGTACAACTCCACCATATGACTAGTGTGCGCTTATGTTGACCACTGGCGAAACTTGGCG 435
Db GAGTACAACTCCACCATATGACTAGTGTGCGCTTATGTTGACCACTGGCGAAACTTGGCG 435
QY 440 CGCATACCGCTGAGAGTCTCTCCACGCGCGCTCCACTCTCTCGGAACTCCG 499
Db CGCATACCGCTGAGAGTCTCTCCACGCGCGCTCCACTCTCTCGGAACTCCG 499
QY 436 CGCATGCTACCTGAGAGTCTCTCCACATAGCTCAAGCTCAAGCGCTTCTATCTGTCGT 495
Db CGCATGCTACCTGAGAGTCTCTCCACATAGCTCAAGCTCAAGCGCTTCTATCTGTCGT 495
QY 500 AGCGACGAGACGAAGCGTCTGATCGAGAGTGTGCTGCGCAAGAACTCGAACGAGAA 559
Db AGCGACGAGACGAAGCGTCTGATCGAGAGTGTGCTGCGCAAGAACTCGAACGAGAA 559
QY 496 AAGATGAGATCCGTGAGTGTCTAC-----TACGTCTCTCCAAACTCGA---GACAT 546
Db AAGATGAGATCCGTGAGTGTCTAC-----TACGTCTCTCCAAACTCGA---GACAT 546
QY 560 GAGTTTGGCGAGTGGAGATTAGTTGATGTTCAACGACTTAATTAACAACAATATG 619
Db GAGTTTGGCGAGTGGAGATTAGTTGATGTTCAACGACTTAATTAACAACAATATG 619
QY 547 GGGTTTGCAGAGTGGAGATGAGACATTTGCTTTTGGAGTTGACTATCAATAACGTTTTT 606
Db GGGTTTGCAGAGTGGAGATGAGACATTTGCTTTTGGAGTTGACTATCAATAACGTTTTT 606
QY 620 AGGATGATATCGGGAAGAGTTTACGGAGAGAGTGAAGAGAGTGAAGAGTTGAGAA 679
Db AGGATGATATCGGGAAGAGTTTACGGAGAGAGTGAAGAGAGTGAAGAGTTGAGAA 679
QY 607 AGAATGATGAGCGGGAACGATTTTATGTTGAGGGCAAGAGTGTGCGCA- 665
Db AGAATGATGAGCGGGAACGATTTTATGTTGAGGGCAAGAGTGTGCGCA- 665
QY 680 GCGAGGAGTTCAGAGAGTGTGACAGAAATGTTGGAATCTATGCGCTTGGCTTAACAAG 739
Db GCGAGGAGTTCAGAGAGTGTGACAGAAATGTTGGAATCTATGCGCTTGGCTTAACAAG 739
QY 666 -----GCAAGTGAACATCTGATGAGAGAGATTGTGACCGCTGGCGCGGAATGCA 720
Db -----GCAAGTGAACATCTGATGAGAGAGATTGTGACCGCTGGCGCGGAATGCA 720
QY 740 GGAGATCATCTTGGCTTCTCAGTGTGTTGATTTTCAAGATGTTGAGAGAGCGCTTAAG 799
Db GGAGATCATCTTGGCTTCTCAGTGTGTTGATTTTCAAGATGTTGAGAGAGCGCTTAAG 799
QY 721 GCCGACTATTTACCAATCATGCTGTTGTTTCA-----AATTTTGAAGAACGCTCAAG 774
Db GCCGACTATTTACCAATCATGCTGTTGTTTCA-----AATTTTGAAGAACGCTCAAG 774
QY 800 AGTATCATGAGTACGATTCCATCTTGAATAGATCTTCTATGAGACCGTGCAGC 859
Db AGTATCATGAGTACGATTCCATCTTGAATAGATCTTCTATGAGACCGTGCAGC 859
QY 775 AATTAGGATTCGTATCGATTAAGTTTGTGAGAGTCTTGTGATGAGAAACGTCGGAT 834
Db AATTAGGATTCGTATCGATTAAGTTTGTGAGAGTCTTGTGATGAGAAACGTCGGAT 834
QY 860 AATGACCGCCAGAAATTCATGATCGATCATCTCTCAACTCGCAAGAGACCCAGCTCAG 919
Db AATGACCGCCAGAAATTCATGATCGATCATCTCTCAACTCGCAAGAGACCCAGCTCAG 919
QY 835 AAGAGAAAGGTTACCATATGATAGTACTTGTCTCTCCAGAAATCTCAGCTGAT 894
Db AAGAGAAAGGTTACCATATGATAGTACTTGTCTCTCCAGAAATCTCAGCTGAT 894
QY 920 TACTACTGACCAAAATCATCAAGGCGCTTCTCTGCGCATGCTTTTGTGGGAACCTGAC 979
Db TACTACTGACCAAAATCATCAAGGCGCTTCTCTGCGCATGCTTTTGTGGGAACCTGAC 979

Db 895 TACTATACAGATGTCACTCTTAAAGGAATATTAATTTGTAATGATAATTTGCTGATCTGAA 954
QY 980 TCATCAACTGGGACTTTAGAGTGGTCAATTAATTAATTTGTAATGATAATTTGCTGATCTGAA 1039
Db 955 ACAATAGCATGGACTTTTAGAATGGCGGATGCTGAATGTTGTAACCTCCAGAGTATTA 1014
QY 1040 AAGAAGGCAAGAGATGAATTTGACACTCAAGTGGGACAAGACCGCTTGTAAATGAGTCA 1099
Db 1015 AAGAAAGCTAGGACCGAAATCGATACGAAATCGTTTTTGACCGGTTAATGATGAAGCT 1074
QY 1100 GACCTTCCAAACTTCCATATCTTAGGAAGATCATCTTGGAGACACTTAGGTTGTACCCC 1159
Db 1075 GATACATAAAATCTGCTTATCTCAATGGATTTGTTGGAGACCTTAGCTTTCACCCG 1134
QY 1160 CGGCGCCCAATTTCTAATACCTCATGTCTTCAGAAAGATATTACAAATTTGAAGGATTTCAAT 1219
Db 1135 GCGCTTCCAAACAAATGTTCCATAGTAGTGGTCAAGAGATGCTGTTGGCAGGATATGAT 1194
QY 1220 ATCCCAACGAGACACAAATTTGATCATTAATGTTGGGCGCATGACAGAGATCCTCAGTTG 1279
Db 1195 GTTCCACGTTGGTTCGATGTTAATGTTGTAACATATGTTCCATGATAGACACCGCTCTATA 1254
QY 1280 TGAATGATGCCATGCTTTTAACTGAGAGTTTGAATGTTGAAGGAGGAGGAGGAGAAAG 1339
Db 1255 TGGGAAGACCCAGAAATGTTCAAGCCAGAGAGGTTCAAAATTTGAAAAGCTTAAATCAAAAG 1314
QY 1340 TTGTGATGCAATTTGGCATGGAAGAGGCTTCCCGAGGAGAACCCATGCTATGCAAGT 1399
Db 1315 TTGTTGCTGTTGGATGGACGACGAGCTGTTCCGGGTATGGCTAGCTCATCGGTA 1374
QY 1400 GTGAGCTTTACTTTGGGATTTGTTGAATTTCAATGTTTGAATGCTGAAACGAGTAAGTAGGAA 1459
Db 1375 GTGAGCTTACTTTGGATGCTGATGTTTCAATGTTTGAATGCTGAAACGAGTAAGTAGGAA 1434
QY 1460 AAGCTTG---ATATGACAGAGAACAAATTTGATGCTTCAAGGTTAATTTCCATTTGGAG 1516
Db 1435 TATGTTGCAATAGTGAAGACAAATGTTGCTTGTGATGCTTCTACGACACCGTTGCTA 1494
QY 1517 GCATGTTGCAAGCTCGCCCACTTGCCTCAATAAATTT 1552
Db 1495 GCCATGTTGAAGCTCGTCCCATTTGTCATGAGATT 1530

RESULT 9

US-09-770-149-144

; Sequence 144, Application US/09770149

; Patent No. US2002005963A1

; GENERAL INFORMATION:

; APPLICANT: Gorlach, Jorn

; APPLICANT: An, Yong-Qiang

; APPLICANT: Hamilton, Carol M.

; APPLICANT: Price, Jennifer L.

; APPLICANT: Raines, Tracy M.

; APPLICANT: Yu, Yang

; APPLICANT: Rameaka, Joshua G.

; APPLICANT: Page, Amy

; APPLICANT: Matthew, Abraham V.

; APPLICANT: Ledford, Brooke L.

; APPLICANT: Woessner, Jeffrey P.

; APPLICANT: Haas, William David

; APPLICANT: Garcia, Carlos A.

; APPLICANT: Kricker, Maja

; APPLICANT: Slader, Ted

; APPLICANT: Davis, Keith R.

; APPLICANT: Allen, Keith

; APPLICANT: Hurban, Patrick

; TITLE OF INVENTION: Expressed Sequences of Arabidopsis

; FILE REFERENCE: 2024 (PARA-013PRV)

; CURRENT APPLICATION NUMBER: US/09/770,149

; CURRENT FILING DATE: 2001-01-26

```

; PRIOR APPLICATION NUMBER: 60/178,506
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FASTSEQ for Windows Version 1.0
; SEQ ID NO 144
; LENGTH: 722
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(722)
; OTHER INFORMATION: n = A,T,C or G
US-09-770-149-144

```

```
Query Match      10.9%; Score 202.8; DB 9; Length 722;
Best Local Similarity 59.6%; Pred. NO. 6.2e-44;
Matches 339; Conservative 0; Mismatches 230; Indels 0
```

RESULT 10

US-10-259-165-565

US-10-239-103-303
; Sequence 565, Application US/10259165
; Publication No. US20030135888A1

; GENERAL INFORMATION:

APPLICANT: Zhu, Tong

; APPLICANT: Wang, Xun

; APPLICANT: Chang, Hur-song

; APPLICANT: Briggs, Stev

APPLICANT: Cooper, Bret

; APPLICANT: Glazebrook, Jane

APPLICANT: Goff, Stephen A.

; APPLICANT: Katagiri, Fumio

```

; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; TITLE OF INVENTION: GENES THAT ARE MODULATED BY POSTTRANSLATIONAL MODIFICATION
; FILE REFERENCE: 70030-NP
; CURRENT APPLICATION NUMBER: US/10/259,165
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/368,327
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 565
; LENGTH: 923
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-259-165-565

```

Query Match 10.4%; Score 192.6; DB 12; Length 923;
Best Local Similarity 55.8%; Pred. No. 4.1e-41;
Matches 388; Conservative 0; Mismatches 304; Indels 3;

RESULT 11
US-10-259-165-433
; Sequence 433, Application US/10259165
; Publication No. US20030135888A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Tong
; APPLICANT: Wang, Xun
; APPLICANT: Chang, Hur-song
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiya
; APPLICANT: Kreps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; TITLE OF INVENTION: GENES THAT ARE MODULATED BY POSTTRANSCRIPTIONAL GENE SILENCING
; FILE REFERENCE: 70030-NP
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US/10/259,165
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 433:
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-259-165-433

Query Match 10.2%; Score 189; DB 12; Length 975;
Best Local Similarity 53.6%; Pred. No. 4e-40;
Matches 471; Conservative 0; Mismatches 390; Indels 18; Gaps 3;

Qy 674 GAGCAGCGGAGGATTCAGAGACTGTGACAGAAATGTTGGNACTCATGGCTGGCT 733
Db 64 GTGGAAGCCAGGAGTTAAGCAGTCTGTCAGAGATCATCCGACATCGGCGGCC 123

Qy 734 AACAAAGGAGATCATCTTCTCAGGTGGTTCGATTTTCAGAAATGT---GGAGAAG 790
Db 124 AACCTGTGGACTATTTCGCGCGCTCGGTGGTTCGACGTGTTCCGCGTCAGGAGGA 183

Qy 791 CGTTAAAGATATCAGTAAGAGTACGATTCATCTTTGAATAAGATCCTTTCAATGA---- 846
Db 184 ATCTCGCGCTGTAAAGCCGAGGAGCGGTTCTCTCGCGCTGTATCGACGCGGCGG 243

Qy 847 -----GRACCTGCCAGCAATGACCGCCAGAAATCCATGATCGATCATCTCTCAA 898
Db 244 CGGAGGTGGACGACGCGGAGGCGGAGAGAGAGATGATCGCGCTGCTCACT 303

Qy 899 CTGCAAGAGACCCAGCTCAGTACTACACTGACCAAAATCATCAAGGCTTGTCTCGGC 958
Db 304 CTGCAGAGACAGAGCGGAGGTGTACACCGATAACATGATCAGCTTAACGCGGAC 363

Qy 959 ATGCTTTTGTGGAACTGATCATCACTGAGCTTTGAGAGTGGTCAATTAATTA 1018
Db 364 TTGTTTCGAGAGGAGCAAGACAACTCTCGACGATCAGAAATGGCGATGTCGCTACT 423

Qy 1019 TTGAATACCCAGAGGTGTTCAAGAGGCAAGAGATGAATTTGGACATCAAGTGGGACAA 1078
Db 424 CTGAACCAACCCGACACACTCAAGAAAGCGAAGCGGAGATCGACGATTCGTCGGAAC 483

Qy 1079 GACCGTTGTTAAATGAGTACAGACTTCCAAAATTTCCATATCTTAGAAGATCATCTT 1138
Db 484 TCTCGCTGATCAGCGCGAGACGATGATCGCTCGCTACCTCCAGTGCATCGTCAGG 543

Qy 1139 GAGACACTTAGGTTGTACCCCGGCCCAATTCATAACCTCAATGTGTCTTCAGAAAT 1198

Db 544 GAGACGCTCCGCTGTACCCGCGCGCGATGCTCTCCCGCACGAGTCTCCCGCGAC 603
Qy 1199 ATTACAAATTGAAGGATTCAATATCCACGAGACACAATTGTGATCATTAATGTTGGGCG 1258
Db 604 TCCAAGGTCCGCGGCTACACATCCGCGCGGCTGATGTTGCTCATCAACGCTACGCC 663
Qy 1259 ATGCAGAGAGATCTCTCAGTTGTGGAATGATGCCACATGCTTTAAACCTGAGAGGTTGAT 1318
Db 664 ATCCACCGTGACCCGCGGTGTGGAGGAGCCGAGAAAGTTTCATGCCGAGAGGTTTCGAG 723
Qy 1319 GTGGAAG---GAGAGGAGAAAGTTGTTAGCAATTTGGCATGGGAGAGGCTTGCCCA 1375
Db 724 GACGCGCGTGCAGCGCAATCTCTTGTATCCGTTTCGGGATGGGAGGCGGAGGTTGCC 783
Qy 1376 GGAGAACCCATGGCTATGCAAAAGTGTCTGCTTTTCTTTGGATTTGTTGATTCATGTTT 1435
Db 784 GCGGAGACGCTGGCGCTGCGCACAGTGGGTTGGTGGGACGCTGATCCAGTGTTC 843
Qy 1436 GACTGGAACGAGTAAAGTGAAGAAAGCTTGTATATGACAGAGACAAATTTGATCACTTG 1495
Db 844 GACTGGAGAGGTTCCAGCGGCTGAGGTTCGACATGACTGAAGTGGCGGCTCACCATC 903
Qy 1496 TCAAGTTAATTCATTTGGNGCCCATGTCRAGGCTCGC 1534
Db 904 CCCAAGGTCGTCCCTTGGAGGCCATGTGTCAGGCGCGC 942

RESULT 12
US-10-259-165-101
; Sequence 101, Application US/10259165
; Publication No. US20030135888A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Tong
; APPLICANT: Wang, Xun
; APPLICANT: Chang, Hur-song
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiya
; APPLICANT: Kreps, Joel
; APPLICANT: Moughamer, Todd
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; TITLE OF INVENTION: GENES THAT ARE MODULATED BY POSTTRANSCRIPTIONAL GENE SILENCING
; FILE REFERENCE: 70030-NP
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/368,327
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentList.pl version 3.0.4 (C) 2001 Syngenta
; SEQ ID NO 101:
; LENGTH: 978
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-259-165-101

Query Match 10.2%; Score 189; DB 12; Length 978;
Best Local Similarity 53.6%; Pred. No. 4e-40;
Matches 471; Conservative 0; Mismatches 390; Indels 18; Gaps 3;

Qy 674 GAGGAAGCGAGGAGTTCAGAGACTGTGACAGAAATGTTGGAACTCATGGCTGGCT 733
Db 64 GTGGAAGCCAGGAGTTAAGCAGGTCGTGACGAGATCATCCGACATCGGCGGCC 123

Qy 734 AACAAAGGAGATCATCTTCTCAGGTGGTTCGATTTTCAGAAATGT---GGAGAAG 790
Db 124 AACCTGTGGACTATTTCGCGCGCTCGGTGGTTCGACGTGTTCCGCGTCAGGAGGA 183

Qy 791 CGTTAAAGATATCAGTAAGAGTACGATTCATCTTTGAATAAGATCCTTTCAATGA---- 846
Db 184 ATCTCGCGCTGTAAAGCCGAGGAGCGGTTCTCTCGCGCTGTATCGACGCGGCGG 243

Qy 847 -----GRACCTGCCAGCAATGACCGCCAGAAATCCATGATCGATCATCTCTCAA 898
Db 244 CGGAGGTGGACGACGCGGAGGCGGAGAGAGAGATGATCGCGCTGCTCACT 303

Qy 899 CTGCAAGAGACCCAGCTCAGTACTACACTGACCAAAATCATCAAGGCTTGTCTCGGC 958
Db 304 CTGCAGAGACAGAGCGGAGGTGTACACCGATAACATGATCAGCTTAACGCGGAC 363

Qy 959 ATGCTTTTGTGGAACTGATCATCACTGAGCTTTGAGAGTGGTCAATTAATTA 1018
Db 364 TTGTTTCGAGAGGAGCAAGACAACTCTCGACGATCAGAAATGGCGATGTCGCTACT 423

Qy 1019 TTGAATACCCAGAGGTGTTCAAGAGGCAAGAGATGAATTTGGACATCAAGTGGGACAA 1078
Db 424 CTGAACCAACCCGACACACTCAAGAAAGCGAAGCGGAGATCGACGATTCGTCGGAAC 483

Qy 1079 GACCGTTGTTAAATGAGTACAGACTTCCAAAATTTCCATATCTTAGAAGATCATCTT 1138
Db 484 TCTCGCTGATCAGCGCGAGACGATGATCGCTCGCTACCTCCAGTGCATCGTCAGG 543

Qy 1139 GAGACACTTAGGTTGTACCCCGGCCCAATTCATAACCTCAATGTGTCTTCAGAAAT 1198

124 AACCTGGGACTTTCGCCGCGCTCCGGTGGTTCGACGTGTTCGGCGTCAGGAGAAG 183
 QY 791 CGTTAAAGAGTATCAGTAAGAGGTACGATTCATCTTGAATAAGATCCTTTCATGA --- 846
 Db 184 ATCTCGCCGCTGTAAGCCGCGGAGGACGCGTTCCTTCGCCGCTGATCGACGCGGAGCG 243
 QY 847 -----GAACCGTGCACCAATTCAGCCGCGGAGGAGGAGGAGGAGGAGGAGGAG 898
 Db 244 CGGAGGCTGACGACGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 303
 QY 899 CTGCAAGAGACCCAGCCCTCAGTACTACACTGACCAAAATCATCAAGGCGCTTCTCTGGCG 958
 Db 304 CTGCAAGAGACAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 363
 QY 959 ATGCTTTTGGTGAAGTACTCATCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1018
 Db 364 TTGTTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 423
 QY 1019 TTGAATCACCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1078
 Db 424 CTGAACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 483
 QY 1079 GACCGGTTGTTAAATCAGTACAGTACGCTTCCAAACTTCCATATCTTAGGAGAGATCATCCTT 1138
 Db 484 TCTCGCTGATACCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 543
 QY 1139 GAGACATTTAGTTGACCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1198
 Db 544 GAGACGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 603
 QY 1199 ATTCAATTGAAGATTCAATATCCACGAGACACAAATTTGATGATCATTAATGTTGGGCG 1258
 Db 604 TGAAGGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 663
 QY 1259 ATGACAGAGATCCTCAGTTGTTGAATGATGCGACATGCTTTAACTCATGATGATGATGATGAT 1318
 Db 664 ATCCAAGTACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 723
 QY 1319 GTGGAAG--GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1375
 Db 724 GACGGGGGTGCGACGCGCAATCTCTTGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 783
 QY 1376 GAGAGACCCATGCTATGAAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1435
 Db 784 GGGAGAGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 843
 QY 1436 GACTGGAAGACGTAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1495
 Db 844 GACTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 903
 QY 1496 TCAAGGTTAAATTCATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1534
 Db 904 CCCAAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 942

RESULT 13

US-10-174-693-221
 ; Sequence 221, Application US/10174693
 ; Publication No. US20030131373A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Boksberg, Leonard N.
 ; TITLE OF INVENTION: Materials and Methods for the
 ; TITLE OF INVENTION: Modification of Plant Lignin Content
 ; FILE REFERENCE: 11000.1003c5
 ; CURRENT APPLICATION NUMBER: US/10/174.693
 ; CURRENT FILING DATE: 2002-06-18
 ; PRIOR APPLICATION NUMBER: US 08/975,316
 ; PRIOR FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: US 08/713,000
 ; PRIOR FILING DATE: 1996-09-11
 ; PRIOR APPLICATION NUMBER: US 09/169,789
 ; PRIOR FILING DATE: 1998-10-09

; PRIOR APPLICATION NUMBER: US 09/615,192
 ; PRIOR FILING DATE: 2000-07-12
 ; NUMBER OF SEQ ID NOS: 407
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 221
 ; LENGTH: 447
 ; TYPE: DNA
 ; ORGANISM: Eucalyptus grandis
 ; US-10-174-693-221

Query Match 9.8%; Score 182.2; DB 15; Length 447;
 Best Local Similarity 64.9%; Pred. No. 1.6e-38;
 Matches 290; Conservative 0; Mismatches 148; Indels 9; Gaps 1;

QY 1038 TGAAGAGGCAAGAGATGAATGGACACTCAAGTGGGCAAGACCGCTTGTAAATGAGT 1097
 Db 1 TAAATGAAGGCCCAAGATGAGATGATCTATGATTTGGCCATGATAGTTTGTAGAAGAT 60
 QY 1098 CAGACCTTCCAAACTTCCATATCTTAGAAGATCATCTTTGAGACACTTTAGGTTGTACC 1157
 Db 61 CGGATGTTTCAAAACTTACCTTACCTTCAGTGCATTAATCTTGAGACCTTTCGACTAAACA 120
 QY 1158 CCCCAGGCGCCCAATCTTAATACCTCATGTGTCTTCCAGAAAGATTAACAATGAAGATTCA 1217
 Db 121 CGACGCGCACCACTTCTCTCCACACGCGTCATCGCTGATTCGACTATAGGAGGATCT 180
 QY 1218 ATATCCCAAGACACAAATTTGATCATTAATGTTGGGGCATGAGAGAGATTCCTCAGT 1277
 Db 181 TCGTCCCAAGGACACTATTTGATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 240
 QY 1278 TGTGAATGATGCCACATGCTTTAAACCTGAGAGGTTTGTATGTAAGAGGAGGAGGAGGAG 1337
 Db 241 TGTGGAGGATGATCAATTTGAGCTTCAAGCCTGAAAGGTTTCAGGCGCAATGGCAGCGAAAGC 300
 QY 1338 A-----GTTGTAGCATTTGGCATGGGAAGAGGCTTGGCCAGGAGAACCCATCG 1388
 Db 301 AACAAAGTACTATTGCTTTGGACATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
 QY 1389 CTATGCAAGTGTGAGCTTTACTTTGGGATGTTGATTCATTAATGATGATGATGATGATGATGAT 1448
 Db 361 CTCATCGGTTCATGGGTGGACGTTGGGCTTGTGATTCAGTGTGATGATGATGATGATGATGAT 420
 QY 1449 TAAGTGAAGGAAAGCTTGCATATGACAG 1475
 Db 421 TAAGCAAGAGAGATTCATGACGCG 447

RESULT 14

US-10-356-153-55
 ; Sequence 55, Application US/10356153
 ; Publication No. US20030166176A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Croteau, Rodney et al.
 ; TITLE OF INVENTION: CYTOCHROME P450 OXYGENASES AND THEIR USES
 ; FILE REFERENCE: 62773
 ; CURRENT APPLICATION NUMBER: US/10/356.153
 ; CURRENT FILING DATE: 2003-01-29
 ; PRIOR APPLICATION NUMBER: US/10/142,231
 ; PRIOR FILING DATE: 2002-05-08
 ; PRIOR APPLICATION NUMBER: 60/165,250
 ; PRIOR FILING DATE: 1999-11-12
 ; NUMBER OF SEQ ID NOS: 95
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 55
 ; LENGTH: 1545
 ; TYPE: DNA
 ; ORGANISM: Taxus cuspidata
 ; US-10-356-153-55

Query Match 9.8%; Score 181.6; DB 12; Length 1545;
 Best Local Similarity 48.5%; Pred. No. 5.4e-38;
 Matches 727; Conservative 0; Mismatches 714; Indels 57; Gaps 6;

Db 340 GTAGGATGACATACAAATCTTCAGTATGGCTCTTACGGTTCCTACTGGCGAAACCTT 399
 Qy 437 CGCGCGATCACCGCCCTCGACGCTCTCTCCAGCGCGGTCCACTCTCTCTCCGGAATC 496
 Db 400 CGGAAATGTGCAGCATCCAGATCTCTCTGCAACAGAAATGGACTCTCTCGACACATC 459
 Qy 497 CGGAGCGACGAGACGAGGCTGTGATGAGAGGTTGGTCTGGCCAGAACTCGAACGAG 556
 Db 460 CGCGTAGAGGAAATTTCT 510
 Qy 557 GAAGAGTTTGGCGGAGTGGAGATTAGTTCAGTCTCAACGACTTAACTTAACTTAACTTAACT 616
 Db 511 CGAGAGACACTCCAGTCAACATGAAGCGAGGCTCTCTGATCTCAGTTAGTATCATC 570
 Qy 617 ATGAGGATGATATCGGGGAAGAGTTTACGGAGGAGAGAGTGAAGTGAAGTGAAGTGAAG 676
 Db 571 CTCCGTATGGTTCGCAACAAAGAAATTAACAGGACCTG---TTTATTCGGAGGAATACGAA 627
 Qy 677 GAAGCGAGGAGTTACAGAGACTGTGACAGAAATGTTGGAATCATGGGCTTGGCTAAC 736
 Db 628 GAAGCGGATCAATTTTAAACAGATGATAAAACAGTCTGTCTTCTTACTTGGAGCAATTTGAG 687
 Qy 737 AAGGAGATCACTTGGCTTTCTCTCAGTGTTCGATTTTTCAGAAATGTTGGAATGTTGGAAGCGCTTA 796
 Db 688 GTTGGAGATTTCTCGCGTTTCTCAAGTGGCTTGAATCTTCAAGGTTTCTAGCTGCTATG 747
 Qy 797 AAGAGTATCAGTAAGAGGTACGATTCATCTTGAATAAGATCTTTCATGAGAACCGTGCC 856
 Db 748 AAAAACTGCAGCAGAAAGAGATGCTTTATGACAGAAATGGTGATTCATCACCGTGAG 807
 Qy 857 AGCAATGACCGCCAGAAATTC-----CATGATCGATCATCTCTCAAACTGCAA 904
 Db 808 AAGAGAGGAGAGTTCGATGCAAAATGCAAGACTTAAATGATGTTCTCTCTCTGCAACA 867
 Qy 905 GAGACCCAGCTCAGTACTAC-----ACTGACCAATCATCAAGGCTTGTCTCTGGCC 958
 Db 868 GACAACCATGAATTCAGTCCGATAGTAACGAGATGTTGTGAAGCCACCGCCCTTACA 927
 Qy 959 ATGCTTTTGGTGGAACTGACTCATCAACTGGGACTTTAGAGTGGTCAATTCATTAATTTA 1018
 Db 928 ATGCTGAAACGAGGTACAGATACATCTCGTGACCATCGAATGGGCAATGGCGGCTCTG 987
 Qy 1019 TTGAATCACCCAGAGGTTTGAAGAGGCAAGAGATGAATTTGACACTCAAGTGGGACAA 1078
 Db 988 ATGACGACCCCTCATATTTTGAAGAAAGCCAGCAGAGCTCGACACGATATCGGACGC 1047
 Qy 1079 GACCGCTTGTAAATGATGATGAGCTTCCAAACTTCCATATCTTAGGAGATCATCCTT 1138
 Db 1048 AGCCGATTTAGAGGAGCAGATCTGACGAGTGAATATTTGAGGCAATTTGGAAT 1107
 Qy 1139 GAGACACTTAGGTTGTACCCCGGCGCCCAATTTCTAATACCTCATGTCTTTCAGAAAT 1198
 Db 1108 GAAACGTTAGGCTATATCCAGCGGACCTCTCTTAGTTCCTCAGAACCATTTAGGAT 1167
 Qy 1199 ATTACAAATTCAGAGATTCATATATCCACAGACACAAATTTGATCATTAATGTTGGGGC 1258
 Db 1168 TGCACTGTTGGAGGTTACCATGTCTCGCAGGAACGCGACTGATTTGATGATGTTGGGCA 1227
 Qy 1259 ATCAGAGAGATCTCAGTTGTGGAATGATGCCACATGCTTTAAACCTGAGAGGT----- 1313
 Db 1228 ATTCACAGACCCGCGAGTGGGAAACCGCCGACCGTGTTCGATCCTGAAACGGTTTGTG 1287
 Qy 1314 -----TTGATGTGAAGGAGAGAGAGAAATTTGGTAGCATTTGGCATG 1357
 Db 1288 AAGAGCGGAAAGAGGTTGACGTAAAGGGCGGAGTTTGAATTTGATTCCGTTGGTTCA 1347
 Qy 1358 GGAAGAGGCTTGCAGAGAGAACCCATGGCTATGCAAGTGTACAGTTTACTTTGGGA 1417
 Db 1348 GGGAGAAGAAATGTGTCCGGGCATGAGTCTGGCATTTGATGTTGTTACGTATACGCTGGGG 1407
 Qy 1418 TTGTTGATTTCAATGTTTGTACTGGAACGAGTAAAGTGAAGAAAGCTTGGATATGACAGAG 1477

Db 1408 AGGCTGCTGCAGAGCTTCGAGTGTCTCTTCCAGAGGTATGATAATTGACATGACGGAA 1467
 Qy 1478 AACAAATTGGATCACCTTGTCAAGGTTAATTCATTGGAGGCCATGTGCAAGGCTCGCC 1535
 Db 1468 GGTTTGGGACTTCACAAATGCCCAAGCAGTTCCTGTTGGAGACCATTTATCAAAACCTCGCC 1525

Search completed: October 21, 2003, 08:02:41
 Job time : 1663 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 21, 2003, 02:43:04 ; Search time 133 Seconds

(without alignments)
6169.408 Million cell updates/sec

Title: US-09-868-546A-1

Perfect score: 1859

Sequence: 1 gaaacacgacgacagacga.....aaataattcttttactttc 1859

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	328.4	17.7	1698	4	US-09-059-769-20
2	182.2	9.8	447	4	US-09-615-192A-221
3	143.4	7.7	494	4	US-09-615-192A-222
4	131.8	7.1	1704	3	US-08-948-564-17
5	128.6	6.9	1821	1	US-08-313-075A-37
6	128.6	6.9	1824	3	US-08-606-505B-1
7	128.6	6.9	1824	3	US-09-616-990-1
8	127.2	6.8	622	2	US-08-975-316-17
9	127.2	6.8	622	4	US-09-615-192A-17
10	120.8	6.5	1812	1	US-07-912-900-28
11	120.8	6.5	1812	1	US-08-285-309-28
12	120.8	6.5	1812	2	US-08-502-046-28
13	118.2	6.4	1755	1	US-07-912-900-29
14	118.2	6.4	1755	1	US-08-285-309-29
15	118.2	6.4	1755	2	US-08-502-046-29
16	117	6.3	2174	3	US-08-606-505B-63
17	117	6.3	2174	3	US-09-616-990-63
18	116.6	6.3	1691	3	US-08-948-564-3
19	114.6	6.2	1929	4	US-09-380-420C-1
20	110.4	5.9	1806	4	US-09-351-229-3
21	109.4	5.9	1893	1	US-08-532-065B-1
22	108.8	5.9	1708	3	US-08-391-677-1
23	108.8	5.9	1761	3	US-09-033-055A-3
24	104	5.6	1824	3	US-08-948-564-13
25	103.2	5.6	1788	3	US-08-948-564-9
26	102.8	5.5	1657	3	US-08-948-564-11
27	101	5.4	1781	4	US-09-499-302A-1

Sequence 2, Appli
Sequence 1, Appli
Sequence 29, Appli
Sequence 24, Appli
Sequence 24, Appli
Sequence 24, Appli
Sequence 201, App
Sequence 67, Appli
Sequence 8, Appli
Sequence 3, Appli
Sequence 69, Appli
Sequence 231, App
Sequence 1, Appli
Sequence 64, Appli
Sequence 64, Appli
Sequence 203, App
Sequence 5, Appli
Sequence 7, Appli

28 99.6 5.4 1634 4 US-09-126-420A-2
29 98.8 5.3 1737 4 US-09-126-420A-1
30 96.8 5.2 1665 1 US-08-313-075A-29
31 93.6 5.0 1666 1 US-07-912-900-24
32 93.6 5.0 1666 1 US-08-285-309-24
33 93.6 5.0 1666 2 US-08-502-046-24
34 91 4.9 993 4 US-09-615-192A-201
35 91 4.9 1665 3 US-09-292-768-67
36 89.4 4.8 1665 3 US-08-881-784-8
37 89.4 4.8 1665 3 US-09-292-768-3
38 89.4 4.8 1665 3 US-09-292-768-69
39 89 4.8 381 4 US-09-615-192A-231
40 88.6 4.8 1838 3 US-08-948-564-1
41 88.2 4.7 1927 3 US-08-606-505B-64
42 88.2 4.7 1927 3 US-09-616-990-64
43 87.4 4.7 432 4 US-09-615-192A-203
44 84.2 4.5 1515 3 US-09-292-768-5
45 83 4.5 1611 3 US-08-948-564-7

ALIGNMENTS

RESULT 1

US-09-059-769-20
; Sequence 20, Application US/09059769
; Patent No. 6329518
; GENERAL INFORMATION:
; APPLICANT: Green, Allan
; APPLICANT: Singh, Surinder
; APPLICANT: Lenman, Marit
; APPLICANT: Stymne, Sten
; TITLE OF INVENTION: Plant Fatty Acid Epoxigenase Genes and Uses
; TITLE OF INVENTION: Therefor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/059,769
FILING DATE: April 14, 1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P06223
FILING DATE: 15-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P06226
FILING DATE: 15-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/043706
FILING DATE: 16-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/050403
FILING DATE: 20-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1698 base pairs
TYPE: nucleic acid


```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 8..1504
US-09-059-769-20

```

Query Match 17.7%; Score 328.4; DB 4; Length 1698;
Best Local Similarity 54.5%; Pred. No. 5.2e-76;
Matches 834: Conservative 0; Mismatches 626; Indels 70;

89	QY	CTTTTCTGGTCTTAATCTCTCGGGTGAAGTTGTTTTCCAAAGCAGAAAATTTGAGAAC	148
44	Db	 CTCTTTCTCTTTCTAATAAAGTTTCAATTTTTTAAACCTCAAAACCCCCCAAAC	103
149	QY	ATACCACGAGTCTCTCTCTTCCCAATAATAGGAACCTTAACTCTCTCGAACAGCCA	208
104	Db	 CTCCCCCATCACACCATCTTTTCCAAATCACCGGCCATCTCCATCTCTCTTAAACAAACA	163
209	QY	ATCCACCGTTTCTTCCAACGATGTCCAAACAGTACGGCAACGTGGTTTCCCTCTGGTTC	268
164	Db	 ATCCACAGAACTCTCCACCAAAATCGCCACCAAGTACGGGACATCTTATTTCTCCGATTC	223
269	QY	GGTTTCAGTCTGGCGTTGTTCATCTCTCTCCAAACAGCATACCAAGAAATGTTTCAACAAA	328
224	Db	 GGAAACAGAAAGTCTTAGTTCATCTCTCTCTCCCGCCGTTACAAGAAATGTTTCACTATA	283
329	QY	CACGAGCTTGCCTTGGCCAAACGGCTACCTTCTCTCGGGAATAATACATCTCTACAAC	388
284	Db	 AAGCATCATTTTCGCTAACCGGCCCAAAATTTCTCGCCGGAAAGCACTTCAATTACAT	343
389	QY	AACACCAACGTAGTCTCTGCTCCACGGGGAGCACTGGCGCAACTCCGCCGATCACC	448
344	Db	 TCCACCACCATGGATTCGGCTCTATGCGGATCACTGGCGTCACTCTCCGACGACTCACA	403
449	QY	GCCTTGACGTCTCTCCACGACGGCGTCCATCTCTCTCGGAATTCGGAGCGACGAG	508
404	Db	 ACAATTGAGCTCTCTCTGCAAAATCGTGTGTTGCCATGTTTTCCGGGTTCCGGCCGATGAA	463
509	QY	ACGAAGCGTCTGATGCAGAGTTGGTCTGTGCCAAGAACTCGAACGAGGAAGTGTTCG	568
464	Db	 AGTACAGCTTTTATCAACAGTTGTTTCCAG-----GAAATCGGGATTCGGGAAG	514
569	QY	CGAGTGGAGATTAAGTTTCGATGTTTCAACGAGCTTAACTTACAAACAATAATGAGGATGATA	628
515	Db	 ATAGTAATTTGACATCGAAACTGATGGAGCTTACACTGTAATACATAATGAGATGCT	574
629	QY	TCGGGGAAGAGTTTACGAGAGGAGTGAATGAAGAACGTTTGAAGAACGAGGAG	688
575	Db	 GCGGAAAACGGTTTACGGGAAAGA-----AGTGAAGGATGAAGAAGGTGAGTTG	625
689	QY	TTCAGACAGACTGTGACAGAAATGTTGGAACTCATGG-----CTTGGCTTAACAGGGAGAT	745
626	Db	 TTGCAGGATCTTATGAAGAAATGGAGCGCTCCGGGGGAATTTCAACGGTGAACAGAGAT	685
746	QY	CACCTTGCCTTCTCAGTGGTTCCGATTTTCAGAAATGTGGAGAAGCGTTAAAGAGTATC	805
686	Db	 TATTTTCCAGTATTCGAGTGGATTGATTACAGGAGTAAAGAAGAAGATCAGGAACTTG	745
806	QY	AGTAAGAGGTACGATTCATCTTGAATAAGATCTTTCATGAGAACCGTGGCAGCAATGAC	865
746	Db	 ATGAAGAAAATGGACGGGTTCTTGCAAAATCTCATTTGATGAACACCGAAACACGACGTTG	805
866	QY	CGCCAGAAAT-----TCCATGATCGATCAT	889
806	Db	 TGGATCAATCAAGTTCGAGCAACTCGGACAAAAAGAGGAACTTGGACACTGGTAGATGTT	865
890	QY	CTCTCTAAACTGCAAGAGACCCAGCCTCAGTACTACACTGACCAAAATCATCAAGGCCTT	949
866	Db	 ATGTTGAATCTTAAAAAAGACACAACTTGACTTCTTACACTGATCTAACTATCAAAAGGTGC	925
950	QY	GCTCTGCCCATGCTTTTTTGGTGGAACTGACTCATCAACTGGACTTTAGAGTGGTCAATTA	1009

[illegible]

```

RESULT 2
US-09-615-192A-221
; Sequence 221, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 221
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-615-192A-221

```

Query Match 9.8%; Score 182.2; DB 4; Length 447;
Best Local Similarity 64.9%; Pred. No. 4.1e-38;
Matches 290; Conservative 0; Mismatches 148; Indels 9; Gaps 1;

QY 1038 TGAGAGGCAAGAGAGTAATTGGACACTCAAGTGGGCAAGAGCCGCTTTGTTAAATGACT 1097
DB 1 TAATGAAGGCCCAAGATGAGATTGATCTATGATTTGGCCATGATAGTTTGTAGAGAAAT 60
QY 1098 CAGACCTTCCAAACCTTCATATCTTAGAAGATCATCTTGAGACACTTAGTGTGAC 1157
DB 61 CGAGTGTTCAAACCTTACCTTACCTTCAAGTGCATTAATCTTGGAGCCCTTCGACTAAACA 120
QY 1158 CCGCGGCCCAATTCCTAATACCTCATGTCTTTCAGAGAGATTAATCAATGAAGGATTC 1217
DB 121 CGAGCGCACCACTTCTCTCCACACGCTCATCGCTGATTCGACTATAGGAGGATACT 180
QY 1218 ATATCCACAGACACAATTTGATCATTAATGTTGGGCGCATGAGAGAGATTCCTCACT 1277
DB 181 TCGTCCACGCGACACTATTTGATGTTGAATGATCGATGGCCATTCACAAAGACCCCTCACT 240
QY 1278 TGTGGAATGATGCCACATCTTTAAACCTGAGAGGTTTGTATGGAAGGAGAGAGAGAAA 1337
DB 241 TGTGGAGGATTCATTTGAGCTTCAAGCTTGAAGGTTTCAGGGCAATGGCAGGAAAGC 300
QY 1338 A-----GTTGGTAGCATTTGGCATGGGAAGAGGGCTTGGCCAGGAGAACCCATCG 1388
DB 301 AACAAAGCTACTATTGCTTTTGGACTGGGACGAGGGCATGCCCTGGTGGCCCTTGG 360
QY 1389 CTATGCAAGTGTACGCTTTACTTTGGGATTTGTTGATTCATGTTTGTGACTGGAACGAG 1448
DB 361 CTCATCGGCTCATGGGGTGGACCTTGGGCTTGTGATTCAGTGTGTTTGTGGAAGAG 420
QY 1449 TAAGTGAAGAAAGCTTGATGATCAGAG 1475
DB 421 TAAGCGAAGAGAGATTGATGACGG 447

RESULT 3

US-09-615-192A-222
; Sequence 222, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Siminszky, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Corbin, Frederick T.
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1996-09-11
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 222
; LENGTH: 494
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-615-192A-222

Query Match 7.7%; Score 143.4; DB 4; Length 494;
Best Local Similarity 57.4%; Pred. No. 5.8e-28;
Matches 283; Conservative 0; Mismatches 201; Indels 9; Gaps 1;

QY 733 TAACAGGGAGATCACATGCTTCTCCCTCAGGTGGTTCGATTTTCAGAAATGTGAGAGCG 792
DB 1 TTACCTTGGGATTTCTTGGCCACTACTAAAGTTGGTGCATTACATGAGTCAAGAGAG 60
QY 793 CTTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 852
DB 61 GGTGTTGAGTGAAGAGAGAAATTCGATGCTTCATTCAGGGCTTGTATCAACGAGACCG 120
QY 853 -----TGCAGAGCAATGACCGCCAGAAATTCATGATCGATCACTCTCAACTGCA 903

DB 121 GAGAGAGAGGGGACCCAGAGAGCTCGCAGACAGATGATGATCTATCTTCTGCATCTACA 180
QY 904 AGAGACCCAGGCTCAGTACTACACTGACCAATCATCAAGGCGCTTGTCTGGCCATGCT 963
DB 181 AGAATCTCAGCGGAGAGACTACTCGGACTCCATGATCAAGGGCTTGTCTGTTTGT 240
QY 964 TTTTGGTGAATGACTCATCACTGGGACTTTAGAGTGGTCAATATCTAATTTATTGAA 1023
DB 241 AGTTGCGGGAACAGACAGCTCATCGTTTACATTAGAATGGATAATGACAAACTTACTAAA 300
QY 1024 TCACCCAGAGGTGTTGAAGAGGCAAGAGATGAATTTGACACTCAAGTGGGACAGACCG 1083
DB 301 CAATCTGAAAGTTAGAGAGGCGCGAAATGAGATTGATTTCTGTTATTTGGCCAGATCG 360
QY 1084 CTTGTTAAATGAGTCAGACCTTCCAAAACTTCCATATCTTAGGAAGATCATCTTTGAGAC 1143
DB 361 TCTGTTAGAGATCGGATGTTTCCGATCTACTTACCTTCAGTGCATCATCTTAGAGAC 420
QY 1144 ACTTAGTGTGACCCCGCGCCCAATTTCTAATACCTCATGCTCTTCCAGAGATATTAC 1203
DB 421 CTTTCGACTAAACACCCAGCGTGCACCTTCTCGTCCCGCAGCATCATCAGCTGATTGCAC 480
QY 1204 AATTGAAGGATTC 1216
DB 481 CATTTGTTGGATAC 493
RESULT 4
US-08-948-564-17
; Sequence 17, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
; APPLICANT: Siminszky, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Corbin, Frederick T.
; TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
; TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia C. Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 6121512th Carolina
; COUNTRY: USA
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400
; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1704 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 38..1564
US-08-948-564-17


```
QY 623 ATGATATCGGGAAGAGGTTTACGGAGAGGAGTGAAGAAAGTTGAGAGCG 682
D 670 ATATGATCGGACAGATGATGCTAAACAAAGAGTATTTG-TAGATAAAGAGTTGAGGTA 728
QY 683 AGGAGAGTTACAGAGAGACTGTGACAGAAATGTTGGAACTCATGGGCTTGGCTAAACAGGA 742
D 729 AATGAATTTAAGACATGTTGTTAGAGTTAATGACATAGCAGGATTTTCAACATTTGTT 788
QY 743 GATCACTTGCCTTCTCAGTGGTTCGATTTTTCAGAAATGTTGGAAGCGCTTAAAGAGT 802
D 789 GATTTATTTCTTGTAGCTGGATGATTTTACAAAGGATAGAAAACGAATGAACGT 848
QY 803 ATCAGTAAGAGTACGATTCATCTTGAATAGATCTTCATGAGAACCGTGCCA----- 857
D 849 TTACATAAGAAGTTTGTATGCTTTATTTGACAAAGATGTTTGTATGAAACAAAGCACTACC 908
QY 858 -----GCAATGACCGCAGAAATTCATGATGATCTCTCTCAAACTGCAAGAGACC 910
D 909 TATGAACGTAAAGGGAACAGATTTCTTGATGTTTGTATGGAATGCGGACAAATTTCT 968
QY 911 CAGCCTCAGTACTACACTGACCAAAATCATCAAGGCTTGTCTCTGCCCATGCTTTTGGT 970
D 969 GAAGGAGAAAGACTCAGTACAAACCAACATCAAGCACATTTTCTGCTGAAATTTGTTACAGCT 1028
QY 971 GGAAGTACTCATCACTGGGACTTTAGAGTGGTCTATCTAATTTATTTGATCACCACCA 1030
D 1029 GGTACGGACACTTCTTCTAGTGAATAGAAATGGGACATTTGCAAGAAATGATGAAGAACCT 1088
QY 1031 GAGGTGTTTCAAGAGCAAGAGATGAATTTGGACACTCAAGTGGGCAAGACCGCTTTGTA 1090
D 1089 GCATTTTGAAGAAAGCAAGCAGAAATGGATCAAGTCAATTTGGAGAAATAGGCGTTTA 1148
QY 1091 AATGAGTACAGCTTCCAAAACCTTCATATCTTAGGAAGATCATCTCTTGAGACACTTAGG 1150
D 1149 CTCGAATCGATATCCAAATCTCCCTTACCTCCGAGCAATTTGCAAGAAACATTTTCA 1208
QY 1151 TTGTACCCCGCCGCCAATCTAATACCTCATGTCTTTCAGAGATATTCAATTGAA 1210
D 1209 AAACACCTTCTACACATTAATCTTCTAGGATCTCGAAGCAACCAATGATGATGAT 1268
QY 1211 GGATTCATATCCACGAGACACAATTTGTATCAATTAATGTTGGGCAATGAGAGAT 1270
D 1269 GGTATTATACCAAAACACTAGGCTTAGTTTAAATGATGGCAATGGAAGAT 1328
QY 1271 CCTCAGTTGTGAATATGCCCATCTTTAAACCTGAGAGGT----- 1313
D 1329 CCCAAGTTTGGGAAATCCACTAGATTTTAAATCCGAAAGATTTCTTGAGTGAAGAAAC 1388
QY 1314 -----TTGATGTGAAGAGGAGGAGAAAGTTGGTAGCATTTGGCAATGGAAGAGG 1366
D 1389 TCCAGATTTGATCTCGAGGGAACGATTTTGAATGATACCAATTTGGTCTGCAAGAGA 1448
QY 1367 GCTTGCCAGGAGAACCCATGGCTATGCAAGGTGTCAGCTTTTACTTTGGGATTTGAT 1426
D 1449 ATTGTGAGGAAACAGAAATGGAAATTTGAATGTTGGAATATATATTAGGAATTTGGTT 1508
QY 1427 CAATGTTTGTACTGGAACAGAGTAAGTGAGG 1457
D 1509 CATTCTTTGATTGGAAATTTACCAAGTGAAG 1539
```

RESULT 6

```
US-08-606-505B-1
; Sequence 1, Application US/08606505B
; Patent No. 6114601
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yasuhiro
; APPLICANT: KIKUCHI, Yasuhiro
; APPLICANT: SHIMADA, Yukihisa
; APPLICANT: OHBAYASHI, Masaya
; APPLICANT: SHIMADA, Ritsuko
; APPLICANT: OKINAKA, Yasushi
; TITLE OF INVENTION: NOVEL PLANT GENES
```

```
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112-3801
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS Ver3.30
; SOFTWARE: PATENT AID Ver1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/606,505B
; FILING DATE: 23-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP44963/92
; FILING DATE: 02-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Perry, Lawrence S.
; REGISTRATION NUMBER: 31865
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-218-2100
; TELEFAX: 212-218-2200
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1824 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Petunia hybrida
; STRAIN: Falcon Blue
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 116 to 1633
; IDENTIFICATION METHOD: by experiment
; US-08-606-505B-1
```

Query Match 6.9%; Score 128.6; DB 3; Length 1824;

Best Local Similarity 49.9%; Pred. No. 8e-24;

Matches 435; Conservative 0; Mismatches 399; Indels 37; Gaps 3;

```
QY 623 ATGATATCGGGAAGAGGTTTACGGAGAGAGTGAAGAAAGTTGAGAGCG 682
D 660 ATATGATCGGACAAAGTATGCTAAAGCAAAAGAGTATTTG-TAGATAAAGGTTGAGGTA 718
QY 683 AGGAGTTTCAGAGAGACTGTGACAGAAATGTTGGAACCTGCTGCTTAAACAGGA 742
D 719 AATGAATTTAAGGACATGTTGTTAGAGTTAATGACAATAGCAGGATTTTCAACATTTGT 778
QY 743 GATCACTTGCCTTCTCAGCTGGTTCGATTTTTCAGAAATGTTGGAAGCGCTTAAAGAGT 802
D 779 GATTTATCTCTGTTTGTAGCTTGGATGATTTTCAAGGATAGAAAACGAATGAACCT 838
QY 803 ATCAGTAAGAGTACGATTTCCATCTTGAATAGATTCCTTTCATGAGAACCGTGCCA----- 857
D 839 TTACATAAGAAGTTTGTATGCTTTATTTGACAAAGATGTTTGTATGAAACAAAGCACTACC 898
QY 858 -----GCAATGACCGCAGAAATTCATGATGATCTCTCTCAAACTGCAAGAGACC 910
D 899 TATGAACGTAAAGGGAACCAAGATTTTCTTGATGTTTGTATGGAATAGGGACAAATTTCT 958
QY 911 CAGCCTCAGTACTACACTGACCAAAATCATCAAGGCTTGTCTCTGCCCATGCTTTTGGT 970
D 959 GAAGGAGAAAGACTCAGTACACCAACCAATCAAGCACATTTTCTGTAATTTGTTTCAGCT 1018
QY 971 GGAAGTACTCATCAACTGGGACTTTAGAGTGGTCTATCTAATTTATTTGATCACCACCA 1030
D 1019 GGTACGGACACTTCTTCTAGTGAATAGAAATGGGCACTTGCAGAAATGATGAAGAACCTT 1078
```

QY 1031 GAGGTGTTGAGAGGCAAGAGATGATTTGGACACTCAAGTGGGACAGACCCCTGTTA 1090
Db 1079 GCCATTTTGAAGAAAGCAGCAGAGAAATGGATCAAGTCAATTTGGAAGAAATAGCGCTTTA 1138
QY 1091 AATGAGTCAGACTTCCAAAGCTTCCATATCTTTAGGAAGATCATCTTGGAGACATTAGG 1150
Db 1139 CTCGAATCCGATATCCCAATCTCCCTTACTTCGGAGCAATTTGCAAGAAACATTTCGA 1198
QY 1151 TTGTACCCCGCGCCCAATTTCTAATACCTCATGTGCTTTTCAAGATATTAACAATTGAA 1210
Db 1199 AAACACCCCTTCTACACCAATTAATCTTCTTAGGATCTCGAAGCAACCATGATGCGAT 1258
QY 1211 GGAATCAATATCCACAGACACAAATTTGATCATTAATGTTGGGCGATGACAGAGAT 1270
Db 1259 GGTATTACATACCAAAAAACACTAGGCTTAGTGTAAACATATGGGCAATTTGGAAGAGAT 1318
QY 1271 CCTCAGTTGTGGATGATGACATGCTTTAAACCTGAGAGGT----- 1313
Db 1319 CCCCAAGTTTGGGAAATCTCACTAGATTTTAATCCCGAAAGATTTCTTGAAGTGAAGAAC 1378
QY 1314 -----TTGATGTGGAAGAGAGGAGGAAAAAGTTGGTAGCATTTGGCATGGGAAGAGG 1366
Db 1379 TCCAAGATTGATCTCTCGAGGAGCAATTTGAATTTGATACCATTTGGTGTGCGACGAGA 1438
QY 1367 GCTTGGCCAGAGAACCCATGGCTATGCAAGTGTGAGCTTTACTTTGGGATTTGAT 1426
Db 1439 ATTGTGTCAGGAAACAAGATGGGAATTTGTAATGGTGAATATATATTAGGAACCTTTGGTT 1498
QY 1427 CAATGTTTCTCACTGGAACGAGTAAGTGAGG 1457
Db 1499 CATTCATTTGTTGGAATTTACCAAGTGAAG 1529

RESULT 7

US-09-616-990-1

; Sequence 1, Application US/09616990

; Patent No. 6232109

; GENERAL INFORMATION:

; APPLICANT: KIKUCHI, Yasuhiro

; KIKUCHI, Yasuhiro

; SHIMADA, Yukihisa

; OHBAYASHI, Masaya

; SHIMADA, Ritsuko

; OKINAKA, Yasushi

; TITLE OF INVENTION: NOVEL PLANT GENES

; NUMBER OF SEQUENCES: 67

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10112-3801

; MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS Ver3.30

; SOFTWARE: PATENT AID Ver1.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/616,990

; FILING DATE: 14-Jul-2000

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP44963/92

; FILING DATE: 02-MAR-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Perry, Lawrence S.

; REGISTRATION NUMBER: 31865

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-218-2100

; TELEFAX: 212-218-2200

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1824 base pairs

;

;

;

;

;

;

;

;

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Petunia hybrida
STRAIN: Falcon Blue
FEATURE:
NAME/KEY: CDS
LOCATION: 116 to 1633
IDENTIFICATION METHOD: by experiment
SEQUENCE DESCRIPTION: SEQ ID NO: 1
US-09-616-990-1

Query Match 6.9%; Score 128.6; DB 3; Length 1824;
Best Local Similarity 49.9%; Pred. No. 8e-24;
Matches 435; Conservative 0; Mismatches 399; Indels 37; Gaps 3;
QY 623 ATGATATCGGGAAGAGGTTTACGGAGAGGAGTGAATGAGATGAAGAACTTTCAGGAGCG 582
Db 660 ATATGATCGGACAGAGTGTCTTAAGCAAAAGAGATTTTG-TAGATAAAGGTTGTGAGGTA 718
QY 683 AGGGAGTTTCAGAGAGACTGTGACAGAAATGTTGGAACTCATCGGCTTGGCTAACAGGGA 742
Db 719 AATGAATTTAAGGACATGTTGTGAGGTTAATGACATAGCAGGGTATTTTCAACATTGT 778
QY 743 GATCACTTGCCTTCTCAGGTGGTTTCGATTTTTCAGAAATGTCGAGAAGCGCTTAAAGAGT 802
Db 779 GATTTTATCTCTTGTAGCTTGGATGGATTTTACAAGGATAGAAAAACGAATGAACGT 838
QY 803 ATCAGTAAAGGTACGATTCATCTTGAATAGATCCTTCATAGAGAACCGTGCCA----- 857
Db 839 TTACATAAAGAGTTTGTGCTTTTATGACAAAGATGTTTGTATGAACAACAACAACTACC 898
QY 858 -----GCAATGACCGCCAGAAATTCATGATCGATCATCTCTCAAACCTGCAAGAGACC 910
Db 899 TATGAACGTAAAGGGAACCAAGATTTTCTGATGTTGTTATGGAATAATGGGCAATTTCT 958
QY 911 CAGCCTCAGTACTACACTGACCAAAATCATCAAAGCGCTTGTCTGTGCCATGCTTTTGGT 970
Db 959 GAAGGAGAAAGACTCAGTACAACCAACATCAAGCACATTTTGTCTGAATTTGTTTACAGCT 1018
QY 971 GGAATGACTCATCACTGGGACTTTAGAGTGGTCTTATCTTAATTTATTTGATCACCCA 1030
Db 1019 GGTACGGACACTTCTTCTAGTCAATAGAAATGGGCATTTGCAGAAATGATGAAGAACCT 1078
QY 1031 GAGGTGTTGAAGAAGCAAGAGATGAATTTGACACTCAAGTGGGCAAGACCGCTTGTGA 1090
Db 1079 GCCATTTTGAAGAAAGCAACAGCAAGATGGATCAAGTCAATTTGGAAGAAATAGCGCTTTA 1138
QY 1091 AATGAGTCAGACTTCCAAAACCTTCCATATCTTAGGAAGATCATCTTGGAGACATTAGG 1150
Db 1139 CTCGAATCCGATATCCCAATCTCCCTTACCTCCGAGCAATTTGCAAGAAACATTTTCA 1198
QY 1151 TTGTACCCCGCGCCCAATTTCTAATACCTCATGTGCTTTCAGAGATATTAACAATTGAA 1210
Db 1199 AAACACCCCTTCTACACCAATTAATCTTCTAGGATCTCGAAGCAACCATGATGCGAT 1258
QY 1211 GGAATCAATATCCCGAGACACAAATTTGTGATCATTAAATGTTGGGCGCATGACAGAGAT 1270
Db 1259 GGTATTACATACCAAAACCACTAGGCTTAGTGTAAATGATATGAGGCAATTTGGAAGAT 1318
QY 1271 CCTCAGTTGTGGAATGATGCCATGCTTTTAAACCTGAGAGGT----- 1313
Db 1319 CCCCAAGTTTGGGAAATCTCACTAGATTTTAAATCCCGAAAGATTTCTTGAAGTGAAGAAC 1378
QY 1314 -----TTGATGTGGAAGAGAGGAGGAAAAAGTTGGTAGCATTTGGCATGGGAAGAGG 1366
Db 1379 TCCAAGATTGATCTCTCGAGGAGCAATTTGAATTTGATACCATTTGGTGTGCGACGAGA 1438
QY 1367 GCTTGGCCAGAGAACCCATGCTATGCAAGTGTGAGCTTTTACTTTGGGATTTGAT 1426
Db 1439 ATTGTGTCAGGAAACAAGATGGGAATTTGTAATGGTGAATATATATTAGGAACCTTTGGTT 1498

SEQUENCE CHARACTERISTICS:

LENGTH: 1812 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-285-309-28

Query Match 6.5%; Score 120.8; DB 1; Length 1812;
Best Local Similarity 50.1%; Pred. No. 8.7e-22;
Matches 437; Conservative 0; Mismatches 397; Indels 38; Gaps 4;

QY 623 ATGATATCGGGAAGAGTTTACCGAGAGGAGAGTGAGATGAAGAAGCTTGAGGAAGCG 682
DB 669 ATATGATCGGACAGATGCTAAGCAAAAGAGTATTG-TAGATAAAGGTTGTGAGGTA 727
QY 683 AGGAGCTTCAGAGAGACTGTGACAGAAATGTTGGAATCATGGCTTGGCTAACAGGGA 742
DB 728 AATGAATTTAAGACATGTTGTAGAGTTAATGACATAGCAGGTTATTCAACATTGGT 787
QY 743 GATCACTTCCTTCTCAGGTGGTTTCAGATTTTCAGATTTGGAGAGCGCTTAAAGAGT 802
DB 788 GATTTTATTCCTTTTACCTGGATGTTTACAGGGATAGAAAAAGAACTGAAACCT 847
QY 803 ATCAGTAAGAGGTACGATTCATCTTGAATAGATCCCTTCATGAGAACCGTGCCA----- 857
DB 848 TTACATAAAGAGTTTGATGCTTTTATTCACAAAGATGTTTGTATGAACACAAAGCAACTACC 907
QY 858 -----GCAATGACCGGAGATTCATGATCATCTCTCAAACTGCAAGAGACC 910
DB 908 TATGAACGTAAGGGGAAACAGATTTTCTTGATGTTGTATGGAATGGGACAAATCT 967
QY 911 CAGCCTCAGTACTACACTGACCAAAATCATCAAAAGGCTTGCTCTGCCCATGCTTTTGT 970
DB 968 GAAGGAGAAGACTCAGTACAACCAACATCARAGCACTTTGCTGATTTGTTTCAGACT 1027
QY 971 GGAACGTACTCACTAGCTGGACTTTAGAGTGTCTAATCTAATTTATGATCAACCA 1030
DB 1028 GGTACGAGCACTTCTTCTAGTGAATAGATGGGCACTTGCAGAAATGATGAGAACCT 1087
QY 1031 GAGGTGTTGAAGAGCAAGAGATGAATGGA-CACTCAAGTGGGCAAGACCGCTTGT 1089
DB 1088 GCCATTTTGAAGAAAGCAAGAGAAATGGATCAAGGTCATTGGAAGAAATAGGCGTTT 1147
QY 1090 AAATGAGTACAGACTTCCAAAATCTTCATATCTTAGAAGATCATCTTGGAGACACTTAG 1149
DB 1148 ACTCGAATCGATATCCAAATCTCCCTTACCTCCGAGCAATTTGCAAGAAACATTTG 1207
QY 1150 GTTGTACCCCGGCCCCCAATTCATACCTCATGTGTCTTCAAGAGATATTAATTTGA 1209
DB 1208 AAAACACCCCTTACACCAATTAATCTTCCAGGATCTCGAACGAACCATGCAATGCGA 1267
QY 1210 AGGATTCATATCCACGAGACAAATTTGATCATTAATGTTGGGCGATGAGAGAGA 1269
DB 1268 TGGTTATCATACCAAAACACTAGGCTTAGTTGTTAATATGAGCAATTTGGAAGAGA 1327
QY 1270 TCCTCAGTTTGGAAATGATCCCATCTTTAAACCTGAGAGT----- 1313
DB 1328 TCCCAAGTTTGGGAAATCCACTAGAGTTTAAATCCCGAAAGATCTTGTAGTGAAGAAA 1387
QY 1314 -----TTGATGTGAAGAGGAGGAGAAAGTTGGTAGCATTTGGCATGGAAGAG 1365
DB 1388 CTCCAAGATTTGATCCTCGAGGAGACGATTTGAAATGATACCAATTTGGTCTCGAGCAAG 1447
QY 1366 GGCTTGCCCGAGGAAACCCATGGCTATGCAAAAGTGCAGCTTTACTTTGGATTTGTGAT 1425
DB 1448 AATTGTGCGAGGAACAAGATGGGAATTTGATGGTGGAAATATATATTAGAACTTTGGT 1507
QY 1426 TCAATGTTTGTACTGGAAACGAGTAAGTGAGG 1457
DB 1508 TCATTCATTTGATGGGAATTTACCAAGTGAAG 1539

RESULT 12

US-08-502-046-28
Sequence 28, Application US/08502046
Patent No. 5861487

GENERAL INFORMATION:

APPLICANT: Holton, Timothy A.
APPLICANT: Cornish, Edwin C.
APPLICANT: Kovacic, Filippa
APPLICANT: Tanaka, Yoshikazu
APPLICANT: Lester, Diane R.
TITLE OF INVENTION: GENETIC SEQUENCES ENCODING A 3,5'-
HYDROXYLASE AND USES

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully, Scott, Murphy & Presser

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: New York

COUNTRY: U.S.A.

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA: US/08/502,046

APPLICATION NUMBER: US/08/502,046

FILING DATE: 14-JUL-1995

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/285,309

FILING DATE: 03-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 8633Z

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 1812 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-502-046-28

Query Match 6.5%; Score 120.8; DB 2; Length 1812;

Best Local Similarity 50.1%; Pred. No. 8.7e-22;

Matches 437; Conservative 0; Mismatches 397; Indels 38; Gaps 4;

QY 623 ATGATATCGGGAAGAGTTTACCGAGAGGAGAGTGAGATGAAGAAGCTTGAGGAAGCG 682
DB 669 ATATGATCGGACAGATGCTAAGCAAAAGAGTATTG-TAGATAAAGGTTGTGAGGTA 727
QY 683 AGGAGCTTCAGAGAGACTGTGACAGAAATGTTGGAATCATGGCTTGGCTAACAGGGA 742
DB 728 AATGAATTTAAGGACATGTTGTAGAGTTAATGCAATAGCAGGTTATTCAACATTGGT 787
QY 743 GATCACTTCCTTCTCAGGTGGTTTCAGATTTTCAGATTTGGAGAGCGCTTAAAGAGT 802
DB 788 GATTTTATTCCTTTTACCTGGATGTTTGTATGGAATGGGATAGAAAAAGAACTGAA 847
QY 803 ATCAGTAAGAGGTACGATTCATCTTGAATAAGATCCTTTCATGAGAACCGTGCCA----- 857
DB 848 TTACATAAAGAGTTTGTATGCTTTTATGACAAAGATGTTTGTATGAGAACCAAGCAACTACC 907
QY 858 -----GCAATGACCGGAGGAAATTCATGATGATCTCTCTCAAACTGCAAGAGACC 910
DB 908 TATGAACGTAAGGGGAAACAGATTTTCTTGATGTTGTATGGAATGGGACAAATCTTCT 967

RESULT 14

US-08-285-309-29
 ; Sequence 29, Application US/08285309
 ; Patent No. 5569832
 ; GENERAL INFORMATION:
 ; APPLICANT: Holton, Timothy A.
 ; APPLICANT: Cornish, Edwina C.
 ; APPLICANT: Kovacic, Filippa
 ; APPLICANT: Tanaka, Yoshikazu
 ; APPLICANT: Lester, Diane R.
 ; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING A 3,5'-
 ; TITLE OF INVENTION: HYDROXYLASE AND USES
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Scully, Scott, Murphy & Presser
 ; STREET: 400 Garden City Plaza
 ; CITY: Garden City
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 11530
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/285,309
 ; FILING DATE: 03-AUG-1994
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: DiGiglio, Frank S.
 ; REGISTRATION NUMBER: 31,346
 ; REFERENCE/DOCKET NUMBER: 86332
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 742-4343
 ; TELEFAX: (516) 742-4366
 ; TELEX: 230 901 SANS UR
 ; INFORMATION FOR SEQ ID NO: 29:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1755 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-285-309-29

Query Match 6.4%; Score 118.2; DB 1; Length 1755;
 Best Local Similarity 49.0%; Pred. No. 4.1e-21;
 Matches 407; Conservative 0; Mismatches 398; Indels 36; Gaps 2;
 QY 663 TGAAGACGTTGAGGAGCGAGGAGTTCAGAGAGACTGTGACAGAAATGTTGCACTCA 722
 DB 613 TAAATAAGGTGTTGAGTAAATGATTTAAGACATGGTGGTAGATTAAAGACACAG 672
 QY 723 TGGCGTTGGCTAAACAGGAGATCACTTGCCTTTCTCAGGTGGTTCGATTTTCAGAATG 782
 DB 673 CAGGGTATTTAACATTGGTGTATTTATTCCTTTAGCTTGGATTTTACAAGGA 732
 QY 783 TGGAGAGCCGTTAAGAGTATCATGTAAGAGTACGATTCATCTTGAATTAAGATCTTC 842
 DB 733 TAGAAAAGGAATGAACCGTTTACATAGAAGTTTGTGCTTTATTGACAAAGATGTTG 792
 QY 843 ATGAGACCGTCCGACGAATGACCGCAGAAATCCATGATCGATCTCTCTCAAACTGC 902
 DB 793 ATGACACAAGCACTAGTATGAACGTAAAGGGAACAGATTTCTTGATTTGTTA 852
 QY 903 AAGAGACCCGCTCAGTACTA-----CACTGACCAATCATCAAAAGGCGCTTG 950
 DB 853 TGGAAAATAGGACAATTTCTGAAGGAGAAAGGCTCAGTACCAACCAATCAAGCACTCT 912
 QY 951 CTCGGCCATGCTTTTGGTGGAACTGACTCATCACTGGGACTTTAGAGTGGTCAATTAT 1010

DB 913 TGCTCAATTTGTTTCAAGCTGGTACAGACACTTCTTCTAGTCAATAGATGGGCACTTG 972
 QY 1011 CTAATTTTGAATCACCAGAGGTTTGAAGAAGGCAAGAGATGAATTTGGACACTCAAG 1070
 DB 973 CAGAGATGATGAAGAACCTTCCATTTTAAAGAAAGCAAGAGGAATGGATCAAGTCA 1032
 QY 1071 TGGGACAAGACCGCTTTGTAATAGTCAAGCTTCCAAAATTTCCATATCTTAGAAGA 1130
 DB 1033 TTGGAACAATAGGCGTCTGCTCGAATCGGATATCCCAATCTCCCTTACCTCCGAGCAA 1092
 QY 1131 TCATCTTTGAGACACTTAGTTGTACCCCGGCCCAATTTCTAATACCTCATGTGCTTT 1190
 DB 1093 TTTGCAAAAGAACATTTGAAAGCACCTTCTACACCAATTAATCTCCCTAGATCTCGA 1152
 QY 1191 CAGAAGATATTACAAATTGAAGGATTTCAATATATCCACGAGACACAATTTGTGATCATTAATG 1250
 DB 1153 ACGAACCATGATTTGCTGATGTTTATACATACCAAAAAACACTAGGCTTAGTGTAAACA 1212
 QY 1251 GTTGGGCAATGAGAGATCTCAGTTGTGGAATGATGCCACATGCTTTTAAACCTGAGA 1310
 DB 1213 TATGGCAATTTGGAAGAGATCCGAAAGTTTGGGAGAACCCACTAGAGTTTTATCTCTGAAA 1272
 QY 1311 GGT-----TTGATGTGGAAGGAGAGAGAAAAGTTGGTAG 1346
 DB 1273 GGTCTTGTAGTGAAGAAACTCGAAGATTTGATCTCGAGGGAACGACTTTGAAATTGATAC 1332
 QY 1347 CATTTGGCATGGAAGAGGCTTCCAGAGGAAACCCATGCTATGCAAAAGTGTGAGCT 1406
 DB 1333 CATTTGGTGTGAGCAAGAAATTTGTGAGGAGCAAGAAATCGTAATGGTGAAT 1392
 QY 1407 TTACTTTGGGATTTGATTTCAATTTTGTGATGCAAAAGGAGTAAGTAGG 1457
 DB 1393 ATATATTAGGAACCTTTGGTCCATTTGATTTGGAATTTACCAAGTGAAG 1443

RESULT 15

US-08-502-046-29
 ; Sequence 29, Application US/08502046
 ; Patent No. 5861487
 ; GENERAL INFORMATION:
 ; APPLICANT: Holton, Timothy A.
 ; APPLICANT: Cornish, Edwina C.
 ; APPLICANT: Kovacic, Filippa
 ; APPLICANT: Tanaka, Yoshikazu
 ; APPLICANT: Lester, Diane R.
 ; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING A 3,5'-
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Scully, Scott, Murphy & Presser
 ; STREET: 400 Garden City Plaza
 ; CITY: Garden City
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 11530
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/502,046
 ; FILING DATE: 14-JUL-1995
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/285,309
 ; FILING DATE: 03-AUG-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: DiGiglio, Frank S.
 ; REGISTRATION NUMBER: 31,346
 ; REFERENCE/DOCKET NUMBER: 86332
 ; TELECOMMUNICATION INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 20:29:01 ; Search time 42 Seconds
(without alignments)
1142.576 Million cell updates/sec

Title: US-09-868-546A-2
Perfect score: 2592
Sequence: 1 LIIWVSAYVFLVFLGVKFP.....RLIPLAMCARPLATKIGI 499

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1287.5	49.7	500	2 C85441	cytochrome P450-11
2	1285	49.6	499	2 T04730	cytochrome P450 ho
3	1241	47.9	497	2 T04731	cytochrome P450 ho
4	1229	47.4	492	2 B85441	cytochrome P450-11
5	1219.5	47.0	495	2 A85441	cytochrome P450-11
6	1212	46.8	543	2 T00513	cytochrome P450 ho
7	1205	46.5	515	2 T00510	probable cytochrom
8	1199.5	46.3	481	2 B96691	probable cytochrom
9	1191.5	46.0	500	2 T52174	cytochrome P450 mo
10	1188.5	45.9	500	2 T04737	cytochrome P450 ho
11	1184.5	45.7	500	2 T52175	cytochrome P450 mo
12	1168.5	45.1	518	2 H85440	cytochrome P450-11
13	1125	43.4	501	2 T04735	cytochrome P450 ho
14	1123	43.3	505	2 T10896	cytochrome P450 (E
15	903.5	34.9	561	2 T49979	cytochrome P450-11
16	888.5	34.3	500	2 T49978	cytochrome P450-11
17	850.5	32.8	513	2 T45624	flavonoid 3'-hydro
18	808.5	31.2	523	2 T04651	cytochrome P450 F1
19	801	30.9	524	2 T04651	cytochrome P450 F1
20	795.5	30.7	543	2 T06533	cytochrome P450 ho
21	784.5	30.3	502	2 T52256	cytochrome P-450LX
22	783	30.2	338	2 T04734	cytochrome P450 ho
23	764.5	29.5	532	2 T05942	cytochrome P450 82
24	762.5	29.4	512	2 T00870	probable cytochrom
25	761.5	29.4	527	2 T07748	probable cytochrom
26	758	29.2	512	2 T00869	probable cytochrom
27	757	29.2	510	2 JC7886	cytochrome P450 92
28	756	29.2	506	2 S38985	flavonoid 3',5'-hy
29	753.5	29.1	498	2 T03634	cytochrome P450 -

RESULT 1

C85441
Cytochrome P450-like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Apr-2001
C:Accession: C85441
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: C85441
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-500 <STO>
A:Cross-references: GB:NC_001268; NID:g7270717; PIDN:CAB80400.1; GSPDB:GN00140
C:Genetics:
A:Gene: At4g37340
A:Map position: 4
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
P:434/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match	49.7%	Score	1287.5	DB 2	Length	500			
Best Local Similarity	49.7%	Pred. No.	1.5e-77						
Matches	247	Conservative	106	Mismatches	129	Indels	15	Gaps	9
Qy	9	VLFVFLGKVFV---QSKRLNTPGPP-PLPIIGNLLEQPTHRFFQRMKQYGN-	63						
Db	6	LIFTFLFISLTLTIIGRIKRNLPFPSPWALPVIGHURLKLPPLHRLVLSVSESLGDA	65						
Qy	64	-VVSLWFGSLAVVISPATVQECTKHVDVALANRLPSLGRKVFYFNNNTVGSCHGEHW	122						
Db	66	PIISLRLGNLRFVVVSSHSLAEECTKNDVLANRFLNSLASKHISYGCYTTVTASYGDHW	125						
Qy	123	RNLRLTALDVLSTORVHFSIGRSDETKRLMORVLAKNNEEFARVVEISSMFNDLTY	182						
Db	126	RNLRLGAVEIFSARLNSFSIRDRTHRLI--ACLSRNS--LEFTKVMKSMFNLT	182						
Qy	183	NNIMRMISGRFPYGESEMKNVVEAREPRETVMLEMLGLANKGPHLPLRFPDPONVE	242						
Db	183	NNIIRMLAGKCYGDBGAB--DDPEAKRVRELTAEAGCCFAGNTADYLPILTW--ITGSE	238						
Qy	243	KRLKSIKRYDSTLANKILHENASND-RQNSMIDHLKLQETOQPYTTDOIIGKLALAML	301						
Db	239	KRIKTIASRLDEFLQGLVDEREGEKEQNTWVDHLLCLQETOPEYTTDOIIGKIMLSLI	298						
Qy	302	FGGTDSTGTLEWLSNLNHNHPEVLKARDELDTQVQDRLNLSLDPKLPYLRKILLET	361						
Db	299	LACTDTSVLTLEWLTISALLNHPQLSKARDEIDNKVGLNRLVEESDLSLHPYLQNI VSES	358						
Qy	362	LRLYPPAPILPHVSSSEDTIEGFNIPRDTIIVINGMQRDPOLMNDATCPKPERFDVE	421						

ALIGNMENTS

```

Db 359 LRLYPASPLLYPHVASECDKGGYHMPRGITMLLTNVAIHDPKIDWDPTSFKPERFEKE 418
Qy 422 GEEKKLAVAGMRGRACPEPMAMQSVFTGLGLLIQCFDWRKRVSEKLDMTTE-NNWITLSR 480
Db 419 GEAQKLLGFLGRRACPGSLAQRLASLTIGSLIQCFEWERIGEEVDMTEGGGVIMPK 478
Qy 481 LIPLAMCKARPLATKI 497
Db 479 AIPLVAMCKARPPVVGKI 495

RESULT 2
T04730
cytochrome P450 homolog F6G17.10 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 28-Jul-2000
C;Accession: T04730
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15382
A;Accession: T04730
A;Molecule type: DNA
A;Residues: 1-499 <BEV>
A;Cross-references: EMBL:AL035601
A;Experimental source: cultivar Columbia; BAC clone F6G17
C;Genetics:
A;Map position: 4
A;Introns: 166/3; 295/3
A;Note: F6G17.10
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homolog
C;Keywords: heme; iron; metalloprotein
F;297-456/Domain: cytochrome P450 homolog <P45>
F;434/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 49.6%; Score 1285; DB 2; Length 499;
Best Local Similarity 50.0%; Pred. No. 2.2e-77;
Matches 251; Conservative 104; Mismatches 129; Indels 18; Gaps 10;

Qy 1 LLVVVSYAVLFLVLFLG-VKVFQSRKLNIPPGPP-PLPIIGNLNLEQPIHRRFFORMS 58
Db 6 LIFTFCFTVLSLIFLIGRIK-----RKL-NLPPSPAWALPVIGHRLKLPPLHVRFLSVS 59
Qy 59 KOYGN--VVSILWFGSLAVVSSPTAYOECPTKHDVALANRLPSLSGKIYFNNTTVGSC 116
Db 60 QSLGDAPITISURLGNLFLVSSSHSIAEECTKNDVILANRQTISTKHISYGNSTVVA 119
Qy 117 SHGHRNRLRITLDVLSQVHSFGIRSDTKRMQRLVLAQNSNEEFARVEISSM 176
Db 120 SYSEHWRNLRIGALEIFSAHLNSFSIRDEIRLLIGLL--RNSS-YGFTKVENKSM 176
Qy 177 FNDLTNNIMRISGRFYGBESSEKMNVEAREFRETVMELMGLANKGDHLPFLRW 236
Db 177 FSDLTENNIRMLAGKCYGDGKE--DDPEAKRVTLIAEAMSSSGPCNAADYIPILTWI 234
Qy 237 DFQNVKELKSIKSYDSILNKILHENASND-RQNSMIDHLLKLOETPOYVTDIIKG 295
Db 235 TYS--ETRIKKLAGLDEFLQCLVDEKREKKEKNTMVDHLLCLOETQPEYMDRIKG 292
Qy 296 LALAMLFQCTSSGTLEWSNLNHPVLEVKARDELDTQVGDRLINESDLKPLPYLR 355
Db 293 TMLSLIAGTDTTAVTLEWALSSLNPNPEVLNKADEIDRMIGVDRLLESDIENLPYLO 352
Qy 356 KIILETLRYPAPILPHVSSEDITIEGFNIPRDTVIINGWGMQDRLNDATCFKP 415
Db 353 NIVSETLRYPAPMLPHVASKCKCVGYDMPRGITMLLTNVAIHDPKIDWDPTSFK 412
Qy 416 ERFQVEGEKKLVAFGMRGRACPEPMAMQSVFTGLGLLIQCFDWRKRVSEKLDMTENN 475
Db 413 ERFEKEGAKMLMPFLGRRACPGSLAQRLVTLISLSLIQCFEWERIGEEVDMTEGPG 472
Qy 476 ITLRLIPLAMCKARPLATKI 497
Db 473 LTMPCARPLEAMCRARDFVGKI 494

```

RESULT 3 T04731

cytochrome P450 homolog F6G17.20 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 28-Jul-2000
C;Accession: T04731
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15382
A;Accession: T04731
A;Molecule type: DNA
A;Residues: 1-497 <BEV>
A;Cross-references: EMBL:AL035601
A;Experimental source: cultivar Columbia; BAC clone F6G17
C;Genetics:

A;Map position: 4
A;Introns: 166/3; 294/3
A;Note: F6G17.20
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homolog
C;Keywords: heme; iron; metalloprotein
F;296-455/Domain: cytochrome P450 homolog <P45>
F;433/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 47.9%; Score 1241; DB 2; Length 497;
Best Local Similarity 47.9%; Pred. No. 1.7e-74;
Matches 237; Conservative 111; Mismatches 133; Indels 14; Gaps 7;

Qy 9 VLFLVLFLGVKVFQSRKLR---NTPGPP-PLPIIGNLNLEQPIHRRFFORMSKOYGN- 63
Db 6 LIFSILFVVLISLTILGKLRKPNLPPSPAWSLFVIGHRLKLPPIHRTFLSLSNNA 65

Qy 64 -VVSILWFGSLAVVSSPTAYOECPTKHDVALANRLPSLSGKIYFNNTTVGSCSGHEW 122
Db 66 PIFSLRNLGNLFLVSSSHSIAEECTKNDVILANRPNFILAKHVAIDYITMIAASYGDHW 125

Qy 123 RNLRRITALDVLSTQVHSFGIRSDTKRMQRLVLAQNSNEEFARVEISSMFDNLT 182
Db 126 RNLRRIGSVEIFSNHRLNSFLSIRKDEIRRLVFL---SRNFSQEFVKVDMKSMLSDLTF 182

Qy 183 NNIMRMISGRFYGBESSEKMNVEAREFRETVMELMGLANKGDHLPFLRWDFONVE 242
Db 183 NNILRWAGKRYIGDGV--DDPEAKRVQLTIADVACAGAGNAVDPVLPVRL--VSDYE 238

Qy 243 KRLKSIKSYDSILNKILHENASND-RQNSMIDHLLKLOETQPYTQDIQIKGLALMLF 302
Db 239 TRVKLAGRLDEFLQCLVDEKREKKEKNTMVDHLLTLOESQPDYFTDRIKGNLMLIL 298

Qy 303 GGTDSSTGTLEWSNLNHPVLEVKARDELDTQVGDRLINESDLKPLPKILLETIL 362
Db 299 AGTDTSAVLEWALSNLNHPVLEVKARDELDTQVGDRLINESDLKPLPKILLETIL 358

Qy 363 RLYPPAPILPHVSSEDITIEGFNIPRDTVIINGWGMQDRLNDATCFKPERFDVEG 422
Db 359 RLYPPAPMLPHVSSDCKVAGYDMPRGITMLLTNVAIHDPKIDWDPTSFKPERKEG 418

Qy 423 EEEKLVAFGMRGRACPEPMAMQSVFTGLGLLIQCFDWRKRVSEKLDMTENNITLSRLI 482
Db 419 EAQKMLPFLGRRACPGSLAQRLVTLISLSLIQCLEWEKIGEE-VDMSEKGVTPKAK 477

Qy 483 PLEAMCKARPLATKI 497
Db 478 PLEAMCRARPPSVVKI 492

RESULT 4 B85441

cytochrome P450-like protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Apr-2001
C;Accession: B85441
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp.

Nature 402, 769-777, 1999

A>Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A;Reference number: A85001; MUID:20083488; PMID:10617198

A;Accession: B85441

A>Status: preliminary

A:Molecule type: DNA

A;Residues: 1-492 <STO>

A;Cross-references: GB:NC_001268; NID:g7270716; PIDN:CAB80399.1; GSPDB:GN00140

C:Genetics:

A;Gene: AT4g37330

A;Map position: 4

C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

C;Keywords: heme; iron; metalloprotein

F;431/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 47.4%; Score 1229; DB 2; Length 492;
Best Local Similarity 47.4%; Pred. No. 1..le-73;
Matches 237; Conservative 107; Mismatches 138; Indels 18; Gaps 7;

QY 2 LVVSVAVLFLVLFLGVKVFQSKRLRNIPGPPPLPIIGNLNLLLEQPIHRRFFQMSK 59
::: :::::
Db 6 ILTSLFSLFTLAIFL----LRSRKLNLPSPALSLPVGHLLHLKPILHRFTLSLK 61
::: :::::
QY 60 QYCN--VSLWGSRLAVVISPTAYOCFTGDVALANRLPSLSGKVIFANNITVGSCS 117
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 62 SIGNARVPHLRGNRLVVYSRSSAEBCFTKNQVYLANRPKFTISKHLGYNATYLLUS 121
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 118 HGHWRNLRIITDALDVLTORVHSFGSGIRSDTKELMQLVLAKNSNEEFARVEISMF 177
:
Db 122 YGDWRNLRIIAVEIFSTRLNSFLYIKDEIRLISHL---SRDSLHGFEVEMKTILL 178
:
QY 178 NDLTYNNIMRMISGRKFGESESMKNVBEAREFRETVTEMLMGLANKGDHPFLRWF 237
:
Db 179 TNLASNTTIRMLAGRYFGEDN-----DDAKLVKLNVSEAVTSAGAGNPIDYLSILRW-- 231
:
QY 238 FQNVKRLKSISKRYDSIINKILHNRAANDQNOMIDHLKLOETQPYTTDOIIGLA 297
:
Db 232 VSSYEKRIRKLNGRFTFLQKLVDKRAEKGETMIDHLLALQIQDPYTYTDVIIKGII 291
:
QY 298 LAMLFGCTSSCTGLEWSLSNLLNHPEVLVKARDDELTCVGDRLNESDLPKLPYLKI 357
:
Db 292 LTUIIAGDTSSVTLEWANSNLLNHPELIKARMEIDEKVGLDRLVDESIVNLSYLSQSI 351
:
QY 358 ILETRLYPAPILIPIHVSESDEITIEGFNIPIRDITVIINGMDRDPOLMNDAFCFKPER 417
:
Db 352 VLETLMYPAVPLLPHLSSECKCVGGYDIPSGTWTLTNAMAHREDPEVWEDEIFKPER 411
:
QY 418 FDVEGEKKLVAFGMGRRACPGEPMAQSVSTLGILLICQFDWKRVSEKLDMTENNWT 477
:
Db 412 FEKEGEAEKLISFGMGRRACPGAGLAHRLINQALGSLVCQFEWERVEDFVDMTEKGAT 471
:
QY 478 LSRLIPLEAMCKARPATKI 497
:
Db 472 LPKAIPURAMCKARSIVDKL 491
: :

RESULT 5

A85441

cytochrome P450-like protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Apr-2001

C;Accession: A85441

R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Col

Nature 402, 769-777, 1999

A>Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana

A;Reference number: A85001; MUID:20083488; PMID:10617198

A;Accession: A85441

A>Status: preliminary

A:Molecule type: DNA

A;Residues: 1-495 <STO>

A;Cross-references: GB:NC_001268; NID:g7270715; PIDN:CAB80398.1; GSPDB:GN00140

C:Genetics:

A;Gene: AT4g37320

A;Map position: 4
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: heme; iron; metalloprotein
F;434/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 47.0%; Score 1219.5; DB 2; Length 495;
Best Local Similarity 47.3%; Pred. No. 4.6e-73;
Matches 232; Conservative 100; Mismatches 147; Indels 11; Gaps 6;

QY 11 FLVFLGKVFV-QSRKLRIIPP-GPP-PLPIIGNLNLEEQPIHRFFORMSKOYGN--VVS 66
DB 11 FLFTISIKULLTSKRKNPPLPPSPAYPLPVIGHLLHLLKQPVRHTFSHSKSLSGNAPIFH 70
QY 67 LWFGRRLAVISSPTAQECCTKHDDVALANPLSLSGKVIFVNNTTVGSCSHEHWRNLR 126
DB 71 LRLGNRLVVISSHSIAEBCFTKNDVVLANPDIIIMAKHVGVNFNTNMIAASYGDHWNR 130
QY 127 RITALDVLSTORVHSPSGIRSDETKRLMORLVLAKNSEEEFARVEISSMFNDLTNNIM 186
DB 131 KIAAVEIPSSHRIISTFSIRKDEIRRLIHL---SRDSLHGFEVELKSLTLNLFANNII 187
QY 187 RMISGKRFYGESBMKNVVEAREPRETVTEMLELMLGLANKGDHLPFLRPFDFONVEKRLK 246
DB 188 MWAGKYGYGTGE--DNDEAKLVRELTAEIMAGAGSNLADYLPsinw-VTNFENQTK 243
QY 247 SISRYDSILNKLIHENRASNDRONSMIDHLKLOETOPOYYTDQIIKGLALAMLGFGTD 306
DB 244 ILGNRLDRVLQKLDVKRAEKKGOTLDHLLSFQETEPEYYTVDVIIKGIILALVLAGTD 303
QY 307 SSTCTLEWSLNLANHPVLKARDDELTOVGQDRLLNESDLPKLPYLARKILLETLRLYP 366
DB 304 TSSVTLEWMSNLNHPHILEKARAEDDKIGSDRLVESDIVNLHYLQNIVSETLRLYP 363
QY 367 PAPILIPIHVSSEDITIEGFNPRTDTIVINGMQORDPOLMNDATCFKEPFDEVEGEKK 426
DB 364 AVPLLPHFSSDECKVAGDYMPRTLTLTNWAMEHDPCLWEPEPFKEPFKEGEARK 423
QY 427 LVAFGMGRACPGFPMAQSVSYFTGLLIQCFDWKRVSEKLDMTENNMITLSRLPLEA 486
DB 424 LMPEFMGRACPGAELGRVLSALGCLITSPEWERVGAELVDMTEGEGITMPKATPLRA 483
QY 487 MCKARPATK 496
DB 484 MCKARAIVGK 493

RESULT 6
T00513
cytochrome P450 homolog At2g23190 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001
C;Accession: T00513; F84621
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.;
submitted to the EMBL Data Library, November 1997
A;Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.
A;Reference number: Z14164
A;Accession: T00513
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-543 <ROU>
A;Cross-references: EMBL:AC002391; NID:g2642427; PID:g2642444
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.L.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; PMID:20083487; PMID:10617197
A;Accession: F84621
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-543 <STO>
A;Cross-references: GB:AEO02093; NID:g2642444; PIDN:AAB87112.1; GSPDB:GN00139

A:Molecule type: DNA
A:Residues: 1-481 <STO>
A:Cross-references: GB:AE005173; NID:g11054542; PIDN:AAG27841.1; GSPDB:GN00141
C:Genetics:
A:Gene: P28G11.4
A:Map position: 1
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F:416/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 46.3%; Score 1199.5; DB 2; Length 481;
Best Local Similarity 47.0%; Pred. No. 9.2e-72;
Matches 237; Conservative 101; Mismatches 123; Indels 43; Gaps 11;

QY 2 LVVSYAVLFLVLGVKVFQSKRLR---NIPGPP-PLPIGNLNLLEQTHIRFFORM 57
Db 8 LILITLITL-----ITLTLIKLKPENLPSPFSLPIIHLRLKLPPLRLFLFI 62

QY 58 SKQYGN--VVSVMFGRSLAVISSPTAYQBCFTKHDVALANRLPSLSGKYIFYNNTVGS 115
Db 63 SKSLNDSPIFSLRGLNKLVPVSSHSIVBECFTKNDVLANR-P----- 105

QY 116 CSHEHWRNLRITLTDVLSQVHSGIRSDTKRLMORLVAKNSNEEFARVEIS 175
Db 106 -----KNLRIGAVEIFENHRLNSFYITRDEIRRLIARLSRPNAS-LEPAKVMNS 157

QY 176 MFNDLTNNIMRMISGRFYGESEEMKNVEAREFRETVMLEMLGLANKGDLPLFLRW 235
Db 158 MSLNLAFFNIIRVMTGKCYGDGA--DDPEAKRVQLIAEAMSCFAGHAADHLPMLRW 215

QY 236 F-PQNVKRLKISIKRYDILNKLHNHNR-ANDRONSMIDHLLKLOETOPQYITDQII 293
Db 216 ITDF---ERRVKIAARLDEFFORLVDKRVAKKENTMIDHLLSLQVSPQYITDHTI 272

QY 294 KGLALAMFGTDSSTGTLEWSLNLNHPVLKARDELDTQVGDRLNLSDELPLKLY 353
Db 273 KGNLSLILAGTDSAVTLEWALSLLNHPVLSKVRDEIDNQIGDLLEESDIINLPY 332

QY 354 LRKIITLRLYPAPILIPHVSSDITIEGFNIPRDTIIVINGMGMORDPOLMNDATCF 413
Db 333 LQNVSETLALYPAGPLLVPHISSDCKVGYDMPGCTMLVNVWAIHRDPRLWDDPASF 392

QY 414 KPRFVDEGEKKLVAFGMGRACPGEPMAQSVFTGLLILOCFDWKRVSEKLDMTEN 473
Db 393 KPERFEKEGTHKLLTGLGRACPGSGLARLVLSLSGLIOCFEWEIRGEBEVDWTEG 452

QY 474 NWITLSRLIPLAEMCKARPLATKI 497
Db 453 GGLTMPRAIPLVAMCRARAFVGKI 476

RESULT 9
T52174
cytochrome P450 monooxygenase [imported] - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 16-Feb-2001
C:Accession: T52174
R:Mizutani, M.; Ward, D.
Plant Mol. Biol. 37, 39-52, 1998
A:Title: Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of cDNAs, differ
A:Reference number: Z14382; MUID:98281573; PMID:9620263
A:Accession: T52174
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-500 <MTZ>
A:Cross-references: EMBL:D78606; PIDN:BA28538.1
C:Genetics:
A:Gene: CYP91A1
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F:301-460/Domain: cytochrome P450 homology <P45>
F:438/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 46.0%; Score 1191.5; DB 2; Length 500;
Best Local Similarity 46.9%; Pred. No. 3.3e-71;
Matches 235; Conservative 100; Mismatches 151; Indels 15; Gaps 7;

QY 4 VVSYAVLFLVLGVKVFQSKRLRNIPGPP-PLPIGNLNLLEQTHIRFFORMSKQY- 61
Db 6 VVLISI-FSLIFLIISFKFLPKQNLPSPGWLPIIHLRLKLPPIHRLRSFSETLD 64

QY 62 ----GNVSVLWFGSRLAVISS-PTAYQBCFTKHDVALANRLPSLSGKYIFYNNTVGS 116
Db 65 HNDGGVMSLRGLSRLVYVSSHVAABECFGKNDVLANRPOVLIIGHGVGYNNAIMAA 124

QY 117 SHGEHWRNLRITLTDVLSQVHSGIRSDTKRLMORLVAKNSNEEFARVEIS 176
Db 125 PYGDHWRNLRILCTIEIFSTRLNCFLYTRDEVRRLISRLSLAGTK--TVVELKPM 181

QY 177 FNDLTNNIMRMISGRFYGESEEMKNVEAREFRETVMLEMLGLANKGDLPLFLRW 236
Db 182 LMDLTFNNIMRMVMTGKRYIGEET--TDEEAKRVKLVADVANTSGNAVDPILRL- 238

QY 237 DFQNVKRLKISIKRYDILNKLHNHNR-ANDRONSMIDHLLKLOETOPQYITDQII 296
Db 239 -FSSYENRVKLGEBETDKFLOGLDDKRGQOETGTMTIDHLLVLOKSDIEYITDQII 297

QY 297 ALAMFGTDSSTGTLEWSLNLNHPVLKARDELDTQVGDRLNLSDELPLKLYRK 356
Db 298 ILIMVIAGTNSAVTLEWALSLLNHPVLSKVRDEIDNRVGLDLRIEADLSLPLKN 357

QY 357 ILETLRLYPAPILIPHVSSDITIEGFNIPRDTIIVINGMGMORDPOLMNDATCF 416
Db 358 IVLETLRLHPATPLLVPHMASEDCIKSYDMPRGITLLVNAWAIHRDPTWDDPSFKPE 417

QY 417 RPDVGEKKLVAFGMGRACPGEPMAQSVFTGLLILOCFDWKRVSEKLDMTENW 476
Db 418 RPEKEEAQKLLAFGLGRACPGSGLAQRIVGLALGSLIQCFEWEIRGEBEVDW 477

QY 477 TSLRLIPLAEMCKARPLATKI 497
Db 478 TVPKAIPKAUKARPLHKI 498

RESULT 10
T04737
cytochrome P450 homolog F6G17.80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 04-Mar-2000
C:Accession: T04737
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15382
A:Accession: T04737
A:Molecule type: DNA
A:Residues: 1-500 <BEV>
A:Cross-references: EMBL:AL035601
A:Experimental source: cultivar Columbia; BAC clone F6G17
C:Genetics:
A:Map position: 4
A:Introns: 295/3
A:Note: F6G17.80
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F:297-460/Domain: cytochrome P450 homology <P45>
F:438/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 45.9%; Score 1188.5; DB 2; Length 500;
Best Local Similarity 45.0%; Pred. No. 5.2e-71;
Matches 225; Conservative 120; Mismatches 142; Indels 13; Gaps 9;

QY 7 YAVLFVLVLGV--KPVFQSKRLRNIPGPPPLPIIIGNLNLLEQTHIRFFORMSKQYGV 64
Db 3 YFILLPLLVISYKLYSKTQRFNLPFGPPSPFVGHLLHLMKPPHRLRLQSYNOYGP 62

QY 65 VSLWFSRLAVISSPTAYQBCFT-KHDVALANRLPSLSGKYIFYNNTVGSCHGEHWR 123

```
Db 63 FSLRFGRRVVVITSPSLAQESFTGQNDIVLSSPLQLTAKYVAYNHTTGTAPYGDHWR 122
Qy 124 NLRITLALDVLSTORVHSFSGIRSDETKRLMORLV-LAKNSNE--EEFARVEISSMFDNLT 181
Db 123 NLRICQEISSLHRLNFQHIRKDEILRMLTRLSRYTQTSNESNDTHLEPLLSDLT 182
Qy 182 YNNIMRMISGRFYGESEEMKNVEAREFRETVTTEMLMGLANKGDLHPLFLRWDFQNV 241
Db 183 FNNIVRMVTKRYGDD--DVNNKEEAEFLFKLVYDIAMYSGANHSADYLPILKFLG-NKF 239
Qy 242 EKRLKSISKRYDSITLNLKILHENRASNDRONSMIDHLLKLQETQPOYTTDQIIKGLALAML 301
Db 240 EKEVKAIGKSMDDILQRLDECRRDKE-GNTVWNHLISLQOQOPEYTTDVIILKGLMWSM 298
Qy 302 FGGTDSSTGTLEWSLSNLLNHPVLEKARDELDTQVGDRLNLSDELPLKPYLRKIILET 361
Db 299 LAGTETSATVLEWAMANLLRNPEVLEKARSEIDEKIGKDLIDESDIAPVLYQNVVSET 358
Qy 362 LRLYPPAPILPHVSSDITIEGFNIPRDTIVIINGWMQORDPOLMNDATCFKPERF-DV 420
Db 359 FRLPVAPFLIPRSPDTDMKIGGYDVRDTIVMVNAWAIHRDPEIWEPEKFNPDYNDG 418
Qy 421 EGEE---KKLVAFGMRACRACPEGMAMQSVFTLGLLIQCFDWKRVSEKLDNTENNWT 477
Db 419 CGSDYYVYKMLPFGNGRRCTPCGAGLQRIIVTLALGSLIQCFEWEVNYKGEEMDSSTGLG 478
Qy 478 LSRILPLEAMCKARPLATKI 497
Db 479 MRKMDPLAMCRPRPIMSKL 498

RESULT 11
T52175
cytochrome P450 monooxygenase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 16-Feb-2001
C;Accession: T52175
R;Mizutani, M.; Ward, E.; Ohta, D.
Plant Mol. Biol. 37, 39-52, 1998
A;Title: Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of cDNAs, differ
A;Reference number: 214382; MUID:98281573; PMID:9620263
A;Accession: T52175
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-500 <MIZ>
A;Cross-references: EMBL:D78607; PIDN:BAA28539.1
C;Genetics:
A;Note: CYP91A2
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: heme; iron; metalloprotein
F;297-460/Domain: cytochrome P450 homology <P45>
F;438/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 45.78; Score 1184.5; DB 2; Length 500;
Best Local Similarity 44.8%; Pred. No. 9.6e-71;
Matches 224; Conservative 120; Mismatches 143; Indels 13; Gaps 9;

Qy 7 YAVLFLVFLGV--KFVFSQSKLRNIPPGPPPLPIGNLNLLEOPIHRFFORMSKOYGVN 64
Db 3 YFILLPLFLVISYKFLYKSTORNLNPPGPPSRPFVGHLMKPPHRLRLQVSNQYGP 62
Qy 65 VSLWFGRLAVISSPTAYQECFT-KHDVALANRLPSLSGKIYFNNNTTGVSGSHGHW 123
Db 63 FSLRFGRRVVVITSPSLAQESFTGQNDIVLSSPLQLTAKYVAYNHTTGTAPYGDHWR 122
Qy 124 NLRITLALDVLSTORVHSFSGIRSDETKRLMORLV-LAKNSNE--EEFARVEISSMFDNLT 181
Db 123 NLRMCQEISSLHRLNFQHIRKDEILRMLTRLSRYTQTSNESNDTHLEPLLSDLT 182
Qy 182 YNNIMRMISGRFYGESEEMKNVEAREFRETVTTEMLMGLANKGDLHPLFLRWDFQNV 241
Db 183 FNNIVRMVTKRYGDD--DVNNKEEAEFLFKLVYDIAMYSGANHSADYLPILKFLG-NKF 239
```

```
Qy 242 EKRLKSISKRYDSITLNLKILHENRASNDRONSMIDHLLKLQETQPOYTTDQIIKGLALAML 301
Db 240 EKEVKAIGKSMDDILQRLDECRRDKE-GNTVWNHLISLQOQOPEYTTDVIILKGLMWSM 298
Qy 302 FGGTDSSTGTLEWSLSNLLNHPVLEKARDELDTQVGDRLNLSDELPLKPYLRKIILET 361
Db 299 LAGTETSATVLEWAMANLLRNPEVLEKARSEIDEKIGKDLIDESDIAPVLYQNVVSET 358
Qy 362 LRLYPPAPILPHVSSDITIEGFNIPRDTIVIINGWMQORDPOLMNDATCFKPERF-DV 420
Db 359 FRLPVAPFLIPRSPDTDMKIGGYDVRDTIVMVNAWAIHRDPEIWEPEKFNPDYNDG 418
Qy 421 EGEE---KKLVAFGMRACRACPEGMAMQSVFTLGLLIQCFDWKRVSEKLDNTENNWT 477
Db 419 CGSDYYVYKMLPFGNGRRCTPCGAGLQRIIVTLALGSLIQCFEWEVNYKGEEMDSSTGLG 478
Qy 478 LSRILPLEAMCKARPLATKI 497
Db 479 MRKMDPLAMCRPRPIMSKL 498

RESULT 12
H85440
cytochrome P450-like protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Apr-2001
C;Accession: H85440
R;anonymous: The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spr
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: H85440
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-518 <STO>
A;Cross-references: GB:NC_001268; NID:g7270714; PIDN:CAB80397.1; GSPDB:GN00140
C;Genetics:
A;Map position: 4
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: heme; iron; metalloprotein
F;446/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 45.1%; Score 1168.5; DB 2; Length 518;
Best Local Similarity 46.7%; Pred. No. 1.2e-69;
Matches 239; Conservative 101; Mismatches 147; Indels 25; Gaps 10;

Qy 1 LLVVVSYAVLFLVFLGVKVFQSKLRNIPPGPP-PLPIGNLNLLEOP-IHRFFORMS 58
Db 5 LLILTLTVAIFTVKIV---LLVTKPNKLNLPSPNICFPPIIGHLLHLLKPLHRLTSLHS 60
Qy 59 KOYGVNVSILWFGSRLAVISSPTAYQECF-TKHDVALANRLPSLSGKIYFNNNTTGVSGS 117
Db 61 HSLGPVFSRLGSRRLAVISSPTAAEECEPLTKNDIVLANRPFINGKYVAYDTSMVTAP 120
Qy 118 HGEWRNLRRITALDVLSTQRVHSPSGIRSDETKRLMORLVLAKNSEEFARVEISSMF 177
Db 121 YGDWRNLRRITALEVFTSTNRLNASAEIRHDEVKMLQKL---HDLVSVERPAKVELRQLL 177
Qy 178 NDLTYNNIMRMISGRFYGESEEMKNVEAREFRETVTTEMLMGLANKGDLHPLFLRWPD 237
Db 178 TGLTLNIMRMVMTGKRRFF-EEDDGGKAGISLRELVAEILSELAADNPADFLPALURWF 236
Qy 238 FQNVKRLKSIKRYDSITLNLKILHENRASNDR---QNSMIDHLLKLQETQPOYTTDQIIK 294
Db 237 YGLVKAIRIGERWDSLLQGLFDEHRAKORLEFNTWIAHLLDSQEKEPHNSDQTK 296
Qy 295 GLALAMLFQGTDSSTGTLEWSLSNLLNHPVLEKARDELDTQV---GQDRILNESDLPKL 351
Db 297 GLILMMVVGTDTSALTVEWAMSNLLNHPQILLETTRQNDIQMETSSSRRLKKEEDLVNM 356
Qy 352 PYLRKIILETLRLYPPAPILPHVSSDITIEGFNIPRDTIVIINGWMQORDPOLMNDAT 411
```

```
Db 357 NYLVNVSEILRLYPVAPLMPVHPVPSDDCVIGFNVPRDTIVLVNLWAIHRDSVWDDPT 416
Qy 412 CFKPERFDVEGEK-----KLVAFGMRGRACGPEPMAMQSVFTLGLLIQCDFDKRWSE 465
Db 417 SFKPERF--EGSDQFGHYNGKMPFGLGRACPGLSLANRVVGLLGSMLQCFEWSGSG 474
Qy 466 EKLDMTENNWITLSRLLPLEAMCKARPLATKI 497
Db 475 GOVDMTEGPGLSLPAEPLVVTCTREMASEL 506

RESULT 13
T04735
cytochrome P450 homolog F6G17.60 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jan-2000
C:Accession: T04735
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15382
A:Accession: T04735
A:Molecule type: DNA
A:Residues: 1-501 <BEV>
A:Cross-references: EMBL:AL035601
A:Experimental source: cultivar Columbia; BAC clone F6G17
C:Genetics:
A:Map position: 4
A:Introns: 292/3
A:Note: F6G17.60
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
F:294-457/Domain: cytochrome P450 homology <P45>

Query Match 43.4%; Score 1125; DB 2; Length 501;
Best Local Similarity 44.7%; Pred. No. 8.3e-67;
Matches 222; Conservative 120; Mismatches 139; Indels 16; Gaps 10;

Qy 9 VFLVLF-LGVKVFQSKRL-NIPGPP-PLPIGNLNLLEQPIHRFFORMSKQYGVV 65
Db 7 ILPLALFLAYKFFFTSKKORYLPPSPSYSLPILGHLHLIKPPVHRLPHRLNHIPIF 66
Qy 66 SLWFGSLVAVVIGSPYQECFT-KHDVALANRLPSLSGKYFYNNNTVGCSSHGEHWN 124
Db 67 YLGLSRAVVIVSSSLARECFQNDVIVSNRPFLTSKIAYNTIATTSYGDHWN 126
Qy 125 LRRITADVLTQVHVSFGIRSDTKRLMQLRVLAKNSNEEFARVEISSMFDLTNN 184
Db 127 LRRICSLVSSKRLANFLHKEEIQRMLTSLRSDARVCKE---VELSSILVLTFFN 182
Qy 185 IMRMISGRFYGESEMKNVVEAREFRETVMLEMLGLANKGDHLPFLWDFQNVK 244
Db 183 IVRMTVKIYYGD--DVSDEAEFLFKLFTFITNSGARHPGSEYLPFMKIFG-GSPEKE 239
Qy 245 LKSLSKRYDLSILNKILHENRASNDRQNSMIDHLLKLOETQPOYYTDOIIKGLALAMFGG 304
Db 240 VKAAKVIDEMLQRLDECKSDK-DGNTMWNHLLSLQDDPEYITDIIIKGLMGLINVAS 298
Qy 305 TDSSTGTLEWSLNLNHPVYLKKARDELDTQVQDRLNLESPLKPLYLKILLETIRL 364
Db 299 SETSALTIEWAMASLANHPKVLQVKLEIDEIIGQDRLIESDIANPLQNVVSETLRL 358
Qy 365 YPAPILIPHVSSEDTIEGNIPTDVIINGMQRDPQLNDATCFKPERPD-VEGE 423
Db 359 HPAAPVLVPRSTAEIDIKIGGYDVPRDTVMVMVNAWAIHRDPLMTPEPRFNPFGNGE 418
Qy 424 E---KLVAFGMRGRACGPEPMAMQSVFTLGLLIQCDFDKRWSEKLDMTENNWITLSR 480
Db 419 KDDVRMLIATGSGRRICPGVGLAHKI VTLALGSLIQCFDKWKVKEIDMSGEGPMARM 478
Qy 481 LIPLEAMCKARPLATKI 497
Db 479 MVPLRALCKTRPINKL 495
```

RESULT 14

T10896

```
cytochrome P450 (EC 1.14.-.-) 81B1c - Jerusalem artichoke
C:Species: Helianthus tuberosus (Jerusalem artichoke)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Feb-2001
C:Accession: T10896
R:Cabello-Hurtado, F.; Bataud, Y.; Salaun, J.; Durst, F.; Pinot, F.; Werck-Reichhart,
J. Biol. Chem. 273, 7260-7267, 1998
A:Title: Cloning, expression in yeast and functional characterization of CYP81B1, a p
A:Reference number: Z17204; MUID:98184826; PMID:9516419
A:Accession: T10896
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-505 <CAB>
A:Cross-references: EMBL:AJ000477; NID:g3059128; PIDN:CAA04116.1; PID:g3059129
A:Experimental source: cv. blanc commun
C:Genetics:
A:Gene: CYP81B1c
C:Function:
A:Description: specifically catalyzes the hydroxylation of medium chain saturated fatt
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: Chromoprotein; fatty acid metabolism; heme; iron; metalloprotein; oxidorec
F:301-463/Domain: cytochrome P450 homology <P45>
F:441/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 43.3%; Score 1123; DB 2; Length 505;
Best Local Similarity 43.5%; Pred. No. 1.1e-66;
Matches 220; Conservative 118; Mismatches 148; Indels 20; Gaps 8;

Qy 2 LVVSVAVLFLVLGLGVKVFQSKRLNIPGPP-PLPIGNLNLLEQPIHRFFORMSKQ 60
Db 6 LUTTTLLLTFTLYLLLR-----RRSSTUPPIFSLPIIGHLYLKPPLRYRTLAKLSAK 60
Qy 61 YGNVSLWFGSLVAVVIGSPYQECFTKHDVALANRLPSLSGKYFYNNNTVGCSSHGE 120
Db 61 HGQILRLQLGFRRLVIVSSPSAAEECTKNDIVFANRPKMLFGKIIGVNVTSLSWSPYGD 120
Qy 121 HWRNLRRITADVLTQVHVSFGIRSDTKRLMQLRVLAKNSNEEFARVEISSMFDNL 180
Db 121 NWRNLRRIASIILSLHRLNEFHDIRVEETRLIIQKLSACNSGS---SQVTWKFSFYEL 177
Qy 181 TYNMIRMISGRFY-GESEMKNVVEAREFRETVMLEMLGLANKGDHLPFLWDFQ 239
Db 178 TLNVMRMISGRFYFGDNPDEL--EEGRFDMDELDTFVLGASNVGYLPLVSLGLGVK 235
Qy 240 NVYKRLKSLSKRYDLSILNKILHENRASN-----DRQNSMIDHLLKLOETQPOYYTDOIIK 294
Db 236 GLEKKILKLEKRDVFFQGLIDQLRKSKGTEDVNVKKTMIELLLSLOETEPEYTTDAMIR 295
Qy 295 GLALAMFGGTSSTGTLEWSLNLNHPVYLKKARDELDTQVQDRLNLESPLKPLYL 354
Db 296 SFVLVLLAAGSDTSAGTMEWNVMLNHPQVLKKAQNEIDSVIGKNCVDESIPNLPYL 355
Qy 355 KXILLETLYLPAPILIPHVSSEDTIEGNIPTDVIINGMQRDPQLNDATCFK 414
Db 356 RCILNLTLYPAGPLVLPVPEASSCCVGVGVNVRGTLIVNQWAIHDDPKVWDEPETFK 415
Qy 415 PERFD-VEGEK--KLVAFGMRGRACGPEPMAMQSVFTLGLLIQCDFDKRWSEKLDMT 471
Db 416 PERFEGLETRDQGLKLLPFGSGRRSCPGEGLAVRMLGWTIGSLIQCFDWEETSEELVDMT 475
Qy 472 ENNWITLSRLIPLEAMCKARPLATKI 497
Db 476 EGPGLTMPKRAIPLVAKCKPRVENTNL 501

RESULT 15
T49979
cytochrome P450-like protein - Arabidopsis thaliana
N:Alternate names: protein F12B17.50
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 16-Feb-2001
```

Search completed: October 20, 2003, 20:34:48
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 20:21:43 ; Search time 25 Seconds
(without alignments)

938.653 Million cell updates/sec

Title: US-09-868-546A-2

Perfect score: 2592

Sequence: 1 LLVVSVYAVLFLVFLGVKFP.....RLIPLEAMCKARPLATKIGI 499

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1655.5	63.9	499	1	C81E_GLYEC
2	1196.5	46.2	502	1	C81D_ARATH
3	1188.5	45.9	500	1	C81F_ARATH
4	850.5	32.8	513	1	F3PH_ARATH
5	817	31.5	544	1	C821_PEA
6	762.5	29.4	512	1	C7C2_ARATH
7	761.5	29.4	527	1	C823_SOYBN
8	758	29.2	512	1	C7C1_ARATH
9	756	29.2	506	1	C751_PETHY
10	753	29.1	471	1	C771_PERAE
11	746.5	28.8	512	1	F3PH_PETHY
12	738.5	28.5	508	1	C753_PETHY
13	738	28.5	510	1	C933_SOYBN
14	735	28.4	522	1	C822_SOYBN
15	731.5	28.2	511	1	C716_NEPPA
16	730	28.2	523	1	C756_CAME
17	729	28.1	510	1	C7DA_SOYBN
18	727	28.0	513	1	C7D2_SOLME
19	719.5	27.8	511	1	C7C4_ARATH
20	715.5	27.6	525	1	C824_SOYBN
21	713	27.5	490	1	C76B_HELTU
22	709	27.4	504	1	C724_ARATH
23	707.5	27.3	512	1	C981_SORBI
24	707	27.3	502	1	C932_SOYBN
25	704.5	27.2	509	1	C931_SOYBN
26	704	27.2	508	1	C983_ARATH
27	702.5	27.1	497	1	C71C_ARATH
28	701.5	27.1	499	1	C719_SOYBN
29	693.5	26.8	490	1	C71P_ARATH
30	687	26.5	516	1	C754_GENTR
31	685.5	26.4	502	1	C718_MENPI
32	684.5	26.4	509	1	C982_SOYBN
33	681.5	26.3	510	1	C755_EUSGR

34 678 26.2 490 1 C71M_ARATH Q9stl1 arabidopsis
35 677.5 26.1 495 1 C71K_ARATH Q9c0k2 arabidopsis
36 676 26.1 523 1 C9B1_GLYEC P93l49 glycyrrhiza
37 673.5 26.0 510 1 C757_EUSGR O4790 eustoma gra
38 672.5 25.9 497 1 C71D_ARATH O49342 arabidopsis
39 671 25.9 490 1 C71J_ARATH Q9t0k0 arabidopsis
40 669.5 25.8 501 1 C72P_ARATH Q9t12 arabidopsis
41 669.5 25.8 505 1 C762_SOLME P37l22 solanum mel
42 669 25.8 505 1 C712_SOLME P37l18 solanum mel
43 668.5 25.8 483 1 C71N_ARATH Q9stl0 arabidopsis
44 666 25.7 504 1 C7D8_SOYBN O81974 glycine max
45 664.5 25.6 486 1 C71O_ARATH Q9stc9 arabidopsis

ALIGNMENTS

RESULT 1

ID C81E_GLYEC STANDARD; PRT; 499 AA.
AC P93l47;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 81E1 (EC 1.14.-.-) (Isoflavone 2'-hydroxylase) (P450
91A4) (CYP 8E-3).
DE CYP81E1 OR CYP91A4.
OS Glycyrrhiza echinata (Licorice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Galegeae; Glycyrrhiza.
OX NCBI_TaxID=46348;
RN [1]
RP SEQUENCE FROM N.A.
RA Akashi T., Aoki T., Kameya N., Nakamura I., Ayabe S.-I.;
RT "Two new cytochrome P450 cDNAs from elicitor-induced Licorice
(Glycyrrhiza echinata L.) cells".
RL (in) Plant Gene Register PR97-167.
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=99009223; PubMed=9790908;
RA Akashi T., Aoki T., Ayabe S.-I.;
RT "CYP81E1, a cytochrome P450 cDNA of licorice (Glycyrrhiza echinata
L.), encodes isoflavone 2'-hydroxylase".
RL Biochem. Biophys. Res. Commun. 251:67-70(1998).
CC -!- FUNCTION: CATALYZES THE HYDROXYLATION OF ISOFLAVONES, DAIDZEIN AND
FORMONONETIN, TO YIELD 2'-HYDROXYISOFLAVONES, 2'-HYDROXYDAIDZEIN,
AND 2'-HYDROXYFORMONONETIN, RESPECTIVELY.
CC -!- PATHWAY: INVOLVED IN THE BIOSYNTHESIS OF ISOFLAVONOID-DERIVED
ANTIMICROBIAL COMPOUNDS OF LEGUMES.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

EMBL; AB001379; BAA22422.1; --
InterPro; IPR001128; Cytochrome_P450.
Pfam; PF00067; p450; 1.
PRINTS; PR00385; P450.
PROSITE; PS00086; CYTOCHROME_P450; 1.
Oxidoreductase; Monooxygenase; Membrane; Heme.
FT METAL 436 436 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 499 AA; 57191 MW; 4D03AFAC818C04DB CRC64;

Query Match 63.9%; Score 1655.5; DB 1; Length 499;
Best Local Similarity 60.6%; Pred. No. 5.5e-88;
Matches 301; Conservative 100; Mismatches 93; Indels 3; Gaps 1;

QY 1 LLVVSVAVLVLFLGVKVFQSRKLRNTPPGPPPLPIIGNLLEQPIHRRFFQMSKQ 60
 Db 3 ILSLSYSVFLALFFIFNFIVIRAKPKNLPGLPGPPSLPIIGNLHLKRLPHRTFKGLSEK 62
 QY 61 YGNVSLWFGSRLAVVSSFTAYQCEKTHDVALANRLPSLSCKYFYFNNYTTVGSCHGE 120
 Db 63 YGHVSLWFGSRLVWVSSASEFCQCTKQNDVLANRPFSLCKYFYFNTTUGSTSYGE 122
 QY 121 HWNRRLRITDALDVLSTQVRSFSGIRSDTKRLMORLVLAKNSNEEFARVEISSMFNDL 180
 Db 123 HWNRRLRITDALDVLNHRINSFSGIRSDTKRLITLADDSSTN--FAEMELSSRLYDM 179
 QY 181 TYNNIMRMISCKPKYGESEKMNVEARBFRTVTWEMLEMLGLANKGDHLPFLRFDQFN 240
 Db 180 TFNNIMRMISCKPKYGECDTSLQEQASQPRDVMVSELLQLSGANNKTFDFMLRFLDFEN 239
 QY 241 VEKRLKSKRYDSILNKILHNENRASNDRONSMDHLLKLQETQPOVYTDQITKGLALAM 300
 Db 240 LEKRLKIDISQDAFLRGLLEHRTKKERANTMIDHLLNLODSQPEYITDQITKGLALAM 299
 QY 301 LFGGTSSTGTLWLSNLNLNHPVLKARDELDTQVQDRLNEDSLPKLPYLKRIILE 360
 Db 300 LLAGTSSAVTLEWMSNLNLNHPVLKVKVDELDTHTVQDRLVDESDDLPLKLYLVKIVNE 359
 QY 361 TRLYPAPILIPHVSSEDTIGFNPRTOTIVINGWQRPQOLWNDATCPKPRFDV 420
 Db 360 TRLYTPAPILPHSTDECNIGYKVPQDTIVLINAWAITHRDPMLWTEATTPKPRFEK 419
 QY 421 EGREKLVAGMGRCAPGPMQVSVFTGLLIQCFDWKRYSEKLDNTENNWTLSR 480
 Db 420 KGELEKLVAGMGRCAPGGLAIRAISMTLALLIQCFLDWKLGKIDLAERDGTLLK 479
 QY 481 LIPLEAMCKARPLATKI 497
 Db 480 LVLPLKAMCKSRPVINKV 496

RESULT 2
 C81D ARATH
 ID C81D ARATH STANDARD; PRT; 502 AA.
 AC Q9FG65; O65789;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cytochrome P450 81D1 (EC 1.14.-.-)
 GN CYP81D1 OR CYP91A1 OR AT5G36220 OR T30G6.3.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxId=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Naruo K., Okumura S., Shino S., Takeuchi C., Wada T.,
 RA Watanabe A., Yanada M., Yasuda M., Sato S., de la Baetide M.,
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney M., Dante M.,
 Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,
 RA Wagner-McPherson C., Wellam A., Yeakum M., Bell M., Dedhia N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vill D., Baker J.,
 RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
 RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
 RA Etian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,

RA Rampsperger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirkse W., Mooijman P., Klein Lankhorst R.,
 RA Weizenecker T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,
 RA Feldpausch M., Lamberth S., Villaroel R., Gielen J., Ardiles W.,
 RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rüd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Fransch P.F.,
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:823-826(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,
 RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 SSP consortium (Salk/Stanford/PGECC).";
 RN Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RP [3]
 RP SEQUENCE OF 3-502 FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=98281573; PubMed=9620263;
 RA Mizutani M., Ward E., Ohta D.,
 RT "Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of
 cDNAs, differential expression, and RFLP mapping of multiple
 cytochromes P450.";
 RL Plant Mol. Biol. 37:39-52(1998).
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB026661; BAB09361.1; --
 CC EMBL; D78606; BAA28538.1; --
 CC EMBL; AY093766; AAM10388.1; --
 CC HSSP; FL4779; 1JF2.
 CC InterPro; IPR001128; Cytochrome_P450.
 CC Pfam; PF00067; P450; 1.
 CC PRINTS; PR00385; P450.
 CC PROSITE; PS00086; CYTOCHROME_P450; 1.
 CC Oxidoreductase; Monooxygenase; Transmembrane; Heme; Multigene family.
 FT TRANSMEM 6 26
 FT METAL 440 440 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT CONFLICT 121 121 T -> A (IN REF. 2).
 SQ SEQUENCE 502 AA; 56723 MW; 1A348A8DD76A41B0 CRC64;
 Query Match 46.2%; Score 1196.5; DB 1; Length 502;
 Best Local Similarity 47.1%; Pred. No. 1.1e-61;
 Matches 236; Conservative 100; Mismatches 150; Indels 15; Gaps 7;
 QY 4 VVSVAVLVLVLFLGVKVFQSRKLRNTPPGPPPLPIIGNLLEQPIHRRFFQMSKQY- 61
 Db 8 VVLSYI-FSLIFLIISFKPLKPKQNLPPSPGWLPIIGHRLKLPPIHRTURSFETLD 66
 QY 62 ----GNVSLWFGSRLAVVSS-FTAYQCEKTHDVALANRLPSLSCKYFYFNNYTTVGS 116
 Db 67 HNDGGGWSRLGSLVYVSSSHKVAEECFCKNDVLANRPPQVVIIGHVGYNTNTWIAA 126
 QY 117 SHGEHWRNLRITDALDVLSTQVRSFSGIRSDTKRLMORLVLAKNSNEEFARVEISSM 176
 Db 127 PYGDHWRNLRITCTIEIFSTRNLNCFLYVYRTDEVRLISRLSLAGTKK---TWELKPM 183
 QY 177 FNDLTNNIMRMISGRFYGESEKMNVEARBFRTVTWEMLEMLGLANKGDHLPFLRWF 236
 Db 184 LMDLTFNNIMRMWTKRYGGEET--TDEEAKRVKLVADVAGNANTSSGNNAVYVPIRL- 240
 QY 237 DFQNVKRLKSKISKRYDSTILNKILHNENRASNDRONSMDHLLKLQETQPOVYTDQITKGL 296
 Db 241 -FSSYENRVKLGEEETDKFQGLIDDKRCQEQETGTTMIDHLLVLQKSDIEYVTDQIIKGI 299

QY 297 ALAMLEGGTDSSTGTTLEMSLNHLNHPVILKARDELDTQVGDRLNLSLDLPKLYLRK 356
 DB 300 ILIMVAGTNTSNTTLEWALSNNLHPDVTISKARDEIDNRVGLDRIEADLSELYLRK 359
 QY 357 IILETLRLVPPAPILIPHYSSSEDTTTEGFNI PRDTTVINGWGMQDRPOLWNDATCFKPE 416
 DB 360 IVLETLRLHPATLPLVPHMASEDECKGSDYMPRGITLLVNAWAIHDPNTWDDPSFKPE 419
 QY 417 RFDVEGEKELVAFMGRRCPGEPNMAQSVSTLGLLIQCQPKRVSEBKDMTENNWI 476
 DB 420 RPEKEEBAOKLAFGLRRACPGSLAQIRVGLGSLGSLIOCFEWRVGVNVDMEKGVGN 479
 QY 477 TSLRLPLEAMCKARPLATKI 497
 DB 480 TVPKAIPLKAICKARPLFKHI 500

RESULT 3
 ID CB1F ARATH STANDARD; PRT; 500 AA.
 AC 065790; Q9SUZ3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Cytochrome P450 81P1 (EC 1.14.-.-)
 GN CYP81P1 OR CYP81A2 OR AT4G37430 OR P6G17.80.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia; TISSUE=Seedling;
 RX MEDLINE=98281573; PubMed=9620263;
 RA Mizutani M., Ward E., Ohta D.
 RA "Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of
 RT cDNAs, differential expression, and RFLP mapping of multiple
 RT cytochromes P450."
 RL Plant Mol. Biol. 37:39-52 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083486; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
 RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidtheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Welljens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weitzenegger T., Bothe G., Ransperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirks W.,
 RA Mociaman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 RA Bernaisel S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
 RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dese C., Fuhs M., Maare A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
 RA Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenot O., Argüel F., Clabaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller K., Schmidt W., Lecharny A., Aubourg S.,
 RA Chedof F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Bielke C.,
 RA Frisman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,

RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stokeling T., Kalicki J., Graves T., Abbott A., Scott K., Johnson D.,
 RA Latreille P., Courtney L., Cloud J., Miller N., Greco T., Kemp K.,
 RA Minx P., Bentley D., Fulton B., Weller N., Pepin K., Hillier L.,
 RA Kraner J., Fulton L., Mardis E., Dantes M., Geisel C., Layman D.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Kameed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RT thaliana."
 RL Nature 402:769-777 (1999).
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D78607; BAA28539.1; -.
 DR EMBL; AL035601; CAB38210.1; -.
 DR PIR; T04737; T04737.
 DR PIR; T52175; T52175.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Membrane; Heme.
 FT METAL 438 438 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT CONFLICT 106 106 A -> V (IN REF. 1).
 FT CONFLICT 127 127 I -> M (IN REF. 1).
 FT CONFLICT 140 140 N -> I (IN REF. 1).
 FT CONFLICT 454 454 S -> T (IN REF. 1).
 SQ SEQUENCE 500 AA; 57555 MW; 0FB453D2070EA2EA CRC64;
 Query Match 45.9%; Score 1188.5; DB 1; Length 500;
 Best Local Similarity 45.0%; Pred. No. 3e-61;
 Matches 225; Conservative 120; Mismatches 142; Indels 13; Gaps 9;
 QY 7 YAVLFVLVLGV--KVFQSRKLRNIPGPPPLPIGNNLLEQPIHRPFQRMKSQYGVN 64
 DB 3 YFILLPLFLVSYKFLYSKTORFNLPGPPSPFVGHLLMKPPIHRLLRQYVNGYPI 62
 QY 65 VSLWFGSLRAVVISPTAYQECFT-KHVDALANRFLPSGKYIFYNNTVVGSGSGHWH 123
 DB 63 FSLRFGSRVWVITPSLAQESFTGQNDIVLSSRPLQLTAKYVYNNHTVTGTPYGDHW 122
 QY 124 NLRITADVLTQVHSPGSRSTETKLMQRLV-LAKNSNE-BEFARVELSSMENDLT 181
 DB 123 NLRICQBSLHSHRLNFQHIRKQDEILRMLRLSRYTQTSNESNDFTHIEPLSLDLT 182
 QY 182 YNNIMRMISGKPYGESEMKVVEAREPREVTVMELMGLANKGDHLPFLRWDFQNV 241
 DB 183 FNNVVMVTGKRYGD--DVNNKEAELEFKLVYDIAMYSNGANHSADYLPILKLF-NKF 239
 QY 242 EKRLKSISKRYDSILNKLILHNRAQNDRNSMIDHLKIQETQPYQYTOIILKGLAML 301
 DB 240 EKEVKAIGKSMDDILQLRLDECRDKE-GNTVWVHLISLQOQOPEYVTVIIGLWMM 298
 QY 302 FGGTSSTGTLWSLSNLLNHPVLKARDELDTQVQDRLLNESDLPKLYLRKILET 361
 DB 299 LAGTETSATVLEWAMNLRNPEVLEKSEIDKIGDRLDESIAVLQNVVSET 358
 QY 362 LRLYPPAPLIPHVSSDITISGFNIPRTIVINGWGMQDRPOLWNDATCFKPERF-DV 420

Db 359 FRLFPVAFPLIPRPTDMKIGGYDVPDRDTVMVNAWAIHRDPEIWEEPKFNPDRYNDG 418
 QY 421 EGEE--KLVAFQMGRRACGPEMAMQSVSFTGLLIQCQFDMKRVSEKLDNTENNWI 477
 Db 419 CGSDYYVYKLPFGNGRTCPGAGLGQRIVTLGSLIQCQFENVKGEEMDSSTGIG 478
 QY 478 LSRLLPLEAMCKARPLATKI 497
 Db 479 MRKMDPLRMCRRPPIKSL 498

RESULT 4
 F3PH ARATH
 ID F3PH ARATH STANDARD; PRT; 513 AA.
 AC Q9SD85;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Flavonoid 3'-monooxygenase (EC 1.14.13.21) (Flavonoid 3'-hydroxylase)
 DE (AtF3'H) (Cytochrome P450 75B1) (TRANSPARENT TESTA 7 protein).
 GN CYP75B1 OR F3'H OR TT7 OR AT5G07990 OR F13G24.190.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia, and cv. Landsberg erecta;
 RX MEDLINE=20483349; PubMed=11030432;
 RA Schoenbohm C., Martens S., Eder C., Forkmann G., Weisshaar B.;
 RT "Identification of the Arabidopsis thaliana flavonoid 3'-hydroxylase
 RT gene and functional expression of the encoded P450 enzyme.";
 RL Biol. Chem. 381:749-753(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Landsberg erecta;
 RX Cordier T.D., Barri-Rewell G., Brugliera F., Cobbett C., Holton T.A.;
 RA "Isolation of a flavonoid 3'-hydroxylase gene corresponding to the Tc7
 RT locus of Arabidopsis thaliana.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Landsberg erecta;
 RX Saslowky D., Winkel-Shirley B.;
 RA "Sequence of flavonoid 3'-hydroxylase (F3'H).";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfling T., Nelson J.,
 RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strowmatt C.,
 RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
 RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
 RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wambutt R., Duesterhoeft A., Stekema W., Pohl T.,
 RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
 RA Rampeger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dikse W., Mooijman P., Klein Lankhorst R.,
 RA Weltzenegger T., Bohe G., Rose M., Hauf J., Berner S., Hempel S.,
 RA Feldpausch M., Lamberth S., Villarroel J., Gielen J., Ardiles W.,

RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Fransz P.F.;
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 408:823-826(2000).
 RN [5]
 RP CHARACTERIZATION.
 RX MEDLINE=21382915; PubMed=11489181;
 RA Saslowky D., Winkel-Shirley B.;
 RT "Localization of flavonoid enzymes in Arabidopsis roots.";
 RL Plant J. 27:37-48(2001).
 CC -!- FUNCTION: Catalyzes the 3'-hydroxylation of the flavonoid B-ring
 CC to the 3',4'-hydroxylated state. Convert naringenin to eriodictyol
 CC and dihydrokaempferol to dihydroquercetin.
 CC -!- CATALYTIC ACTIVITY: A flavonoid + NADPH + O(2) = 3'-
 CC hydroxyflavonoid + NADP(+) + H(2)O.
 CC -!- PATHWAY: Flavonoid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: High expression in siliques and to a lower
 CC extent in stems, flowers and senescing leaves.
 CC -!- INDUCTION: By UV light treatment.
 CC -!- MISCELLANEOUS: May act as a membrane anchor for localization of
 CC other, soluble, flavonoid enzymes to the endoplasmic reticulum.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF271651; AAG16746.1; -
 CC EMBL; AF271650; AAG16745.1; -
 CC EMBL; AF155171; AAF73253.1; -
 CC EMBL; AF241646; AAF60189.1; -
 CC EMBL; AF241643; AAF60189.1; JOINED.
 CC EMBL; AF241644; AAF60189.1; JOINED.
 CC EMBL; AF241645; AAF60189.1; JOINED.
 CC EMBL; AL133421; CAB62611.1; -
 CC PIR; T45624; T45624.
 CC InterPro; IPR001128; Cytochrome_P450.
 CC Pfam; PF00067; P450; 1.
 CC PRINTS; PR00385; P450.
 CC PROSITE; PS00086; CYTOCHROME P450; 1.
 CC Flavonoid biosynthesis; Oxidoreductase; Monooxygenase;
 CC Endoplasmic reticulum; Heme; NADP; Signal-anchor; Transmembrane.
 CC TRANSMEM 1 21
 CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 CC (POTENTIAL).
 CC DOMAIN 22 513
 CC CYTOPLASMIC (POTENTIAL).
 CC METAL 445 445
 CC IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 CC SEQUENCE 513 AA; 56786 MW; C0C740FBE4559C40 CRC64;
 CC
 CC Query Match 32.8%; Score 850.5; DB 1; Length 513;
 CC Best Local Similarity 35.5%; Pred. No. 7e-42;
 CC Matches 182; Conservative 100; Mismatches 193; Indels 37; Gaps 8;
 CC
 CC QY 2 LVVSVAVLVFLVFLGVKVFQSKLRN---IPGPGPPPLPIIGNLLEQPIHRRFQRM 57
 CC Db 6 LTLLATVFLIL-----RIFSHRRNRSHNNRUPPGNPWPIIGNLPHMGTKPHRTLSAM 60
 CC QY 58 SKQYGNVSVLWFGSRVLAVVSSPTAYQECFTKHDVALANRLPSLSGKYIFYNNTTVGSCS 117
 CC Db 61 VTTYGPIILHLRLGFDVVVAASKVAEQFLKIHDFANFASPPPSGAKHMAYNQDLVFP 120
 CC QY 118 HGBHWNRLRITADVLSQTVHSFGIRSDFTKRLMQRLVLAKNNEEFARVEISSMF 177
 CC Db 121 YGHRWRLLARKISSVHLFSKALEDFKVRQEEVGTTLRELVRVGTK-----PVLGLQVL 174
 CC QY 178 NDLTYNNIMRMISGKFPYGESEMKVVEAREPREVTVMLEMLGLANKGDHLPELRWFD 237
 CC Db 175 NMCVVALGEMIGRRRLFGADADHK----ADEFRSMVTMMALAGVFNIGDFVFSLOWLD 230


```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AC003680; AAC06158.1; -.
DR PIR; T00870; T00870.
DR HSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR KW Oxidoreductase; Monooxygenase; Heme; Transmembrane; Multigene family.
FT TRANSMEM 3 23 POTENTIAL.
FT METAL 451 451 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 512 AA; 57256 MW; 4PF5A1BE1C24C798 CRC64;

Query Match 29.4%; Score 762.5; DB 1; Length 512;
Best Local Similarity 34.2%; Pred. No. 7.6e-37;
Matches 175; Conservative 110; Mismatches 187; Indels 39; Gaps 12;

Qy 1 LLVVVAVLFLVLFLGVKVFQGRKLRNIPGPPPLPIIGNLLQPIHRPFQRMKQ 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
9 LFLFLCFLSLPFIFFITTRPSRKVVSPGPPPLPIIGNIHLVGRNPHHSFADLSKT 68
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61 YGVVSLWFGSLAVISSPAYQECFTKHDVALANRLPSLSGKYIFYNNTV---GSC 116
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
69 YGPIMSLKFGSLNTVVVTSPEAREVLTQDILSSRTPTNSIRSNHDKVSVVWLPSS 128
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 SHGHEWRNLRITDALDVLSTQRVHSFGIRDETKRLMQRLVLAKNSNEEFARVEISSM 176
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
129 S---RWLLRLKLSATQLSPQRIEATKTLRENKVELVS--FMSESSEREE--AVDISR- 180
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
177 FNDITYNNIMWISGKRFYGESEKMNVEARE--FRETVMELMLGLANKGDHLPFLR 234
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 ---ATFTALNIISNLF--SVDLGNYSNKGVFQDTVIGVMEAVGNPDAAFFPFLG 234
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
235 WFDQNVKRLKLSIKRVDLSL-----NKLHENRAANDRONSMIDHLLKQI--TQ 284
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
235 FLDLQGNKTKLAKSERLKFVFRGFDIAKAEKSLRDTNSKQVREDFVDVLLDLTSGDE 294
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
285 PQYTDQIIGLALAMLFGGTDSGTGLEWSLSNLLHNPVLEKARDELDTQVQDRLN 344
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
295 AELNTDIVH--LLLDLFCAGTDTNSSVTEWMAELLRNPTWKAQAEIDCVIGKGVVE 353
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
345 ESDLPKLPYLKRIILETLRLYPPAPILPHVSSDITIEGNIIPRDTIVIINGMQRDP 404
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
354 ESDISALPYLOAVVKETFLHPAAPLLVPRKAESDVEVLGFWVPKDTQVFVNVWAIGRP 413
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
405 QLWNDATCFKPERE-----DVEGEKKLVAFGMGRACPGEPMAQSVSFTLGLLIQCFD 459
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
414 NVWENSRFKPERFLGKDIDLRGRDYELTPFGAGRRICPGPLAVKTVPLMLASLLYSFD 473
Qy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
460 WK---RVSEEEKLDMTENNWTLSRLIPLEAM 487
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
474 WKLPNGVSGEDLMDMTFTGLLHKTNPLHAV 504

RESULT 7
C823 SOYBN
ID C823 SOYBN STANDARD; PRT; 527 AA.
AC O49858;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 82A3 (EC 1.14.-.-) (P450 CP6).
GN CYP82A3.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Harosoy 63;

```

```

RX MEDLINE=98311068; PubMed=9648734;
RA Schopfer C.R., Ebel J.;
RT "Identification of elicitor-induced cytochrome P450s of soybean
RT (Glycine max L.) using differential display of mRNA.";
RL Mol. Gen. Genet. 258:315-322(1998).
CC -I- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y10982; CAA71876.1; -.
DR PIR; T07748; T07748.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR KW Oxidoreductase; Monooxygenase; Membrane; Heme.
FT METAL 464 464 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 527 AA; 59822 MW; CA3429E87B202210 CRC64;

Query Match 29.4%; Score 761.5; DB 1; Length 527;
Best Local Similarity 34.6%; Pred. No. 9e-37;
Matches 175; Conservative 103; Mismatches 199; Indels 29; Gaps 7;

Qy 8 AVLFLVLFLGVKVFQGRKLRNIPGPPPLPIIGNLL--EQPIHRFFORMSKQGVNV 65
Db 20 SLIFLCFLYLRK-----NSRGDAPVVGAWPIGLHLSLLNGSQTPHKVLCALADKTGPLF 75
Qy 66 SLWFGSLAVISSPAYQECFTKHDVALANRLPSLSGKYIFYNNTVSGSCGHEWRNL 125
Db 76 TIKLGMKPALVLSNWEWSKELETTNDLAVSSRPKLAVAVMSYNOAFVGLAPYGYREL 135
Qy 126 RRITADVLSSTORVHFSFGIRDETKRLMQRLV---LAKNSNEEFARVEISSMFDLTY 182
Db 136 RKIVTPEFLSNRIEQRNHRVSEVTSIKELFDIWSNGNKNESRYTLVDIKOMLAYLTF 195
Qy 183 NNIMRMISGKRFYGESEKMNVEAREFRETVMELMLGLANKGDHLPFLRWFDFQNV 242
Db 196 NMVVRVVGKRYFG-VNHVEGKDKAQRFMKNTREFNMLMGTTFTVADGVPCLRMLDLGGHE 254
Qy 243 KRLKLSIKRYDSTILNKLHENR-----ASNDRONSMIDHLLKLOETOPQYYTDQI 292
Db 255 KAMNAKEVDKLLSEWSEHQRKLLGENVESDRDFMDVMISALNGAQ--ICAFDADTI 312
Qy 293 IKGLALAMLFGGTDSGTGLEWSLSNLLHNPVLEKARDELDTQVQDRLNLSDLPKLP 352
Db 313 CKATSLEILGGTDSSTAVTLTWALSLLRNPLALGAKAEIDMOIGKDEYIRESDISKLV 372
Qy 353 YLRKIILETLRLYPPAPILPHVSSDITIEGNIIPRDTIVIINGMQRDPOLWMDATC 412
Db 373 YLOAIVKETLRLYPPAPFPSPREFTENCILGGVHIKKKFTLHNLKIHEDPSPVSDPLE 432
Qy 413 FKPERF-----DVEGEKKLVAFGMGRACPGEPMAQSVSFTLGLLIQCFDVKRVS 465
Db 433 FKPERFLTHKDVLDLGRHNFELLPFGSGRRVCAGMSGLNMMVHFTLANLHLSFDILNPSA 492
Qy 466 EKLDMTENNWTLSRLIPLEAMCKAR 491
Db 493 EPVDMTEFFGFTNTKATPLEILVKPR 518

RESULT 8
C7C1 ARATH
ID C7C1 ARATH STANDARD; PRT; 512 AA.
AC O64636; O65783;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

FT	TRANSMEM	3	23	POTENTIAL.
FT	METAL	450	450	IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT	VARSPLIC	307	322	DMFTAGTDTSSSTLEW -> VSLTLQIINIYIME (in isoform Short).
FT				/FTID=VSP_000619.
FT	VARSPLIC	323	512	Missing (in isoform Short).
FT				/FTID=VSP_000620.
FT	CONFLICT	323	323	A -> P (IN REF. 1).
FT	SEQUENCE	512 AA;	56800 MW;	314F9DEBF8C9B54B CRC64;
SQ				
 Query Match				
Best Local Similarity 34.8%; Pred. No. 1.4e-36; Gaps 10				
Matches 176; Conservative 107; Mismatches 193; Indels 30; Gaps 10				
QY	1	LIVVVSYAVLFLVLGLGVKFQSGKRLNIPGPPPLPIICNLMLLEQPIHREFORMSKQ	60	
DB	9	LLLLFCFILSCFLIFTITRSGRISRGATALPGGPRLPIIGNIHLLVGKHPRFAELSKT	68	
QY	61	YGNVVSLMFGSRLAVISSPAYOCECTKHUVALANRPLPSLSKYIFYNNTT-VGSCSHG	119	
DB	69	YGVNMSKLGLSINTVIASPEAREVELRTHDQILSARSPTNAVRSIHQDASLVMLPSSS	128	
QY	120	EHRNLRIRITALDVLSRTQRVHSFGIRSDETCKLMQRLVKAKSNEEEFARVEISSMFND	179	
DB	129	ARWLRLRLSVTQLLSPORIEATKALRNKKVKELVS--FISESDREE--SVDISRVAFI	184	
QY	180	LTYNIMEMI SGKRP--YGESESMKNVEAREPREFTVTMELEMGLANKGDHLFFLRWFD	233	
DB	185	TTLNIISNILSFVDLGSYNAKASINGVQ-----DTVISYMDAAGTPDAANYFPFLAFLD	238	
QY	238	FQNWKEKRLKSISKRYDSILN-----KILHENRASNR---QNSMIDHLLKLOEQPOYYT	289	
DB	239	LQGNVKTEKVCTERLVRVFRGFIDAKIAEKSSQNPKDVKSNDFVDNLLDYKGDELSI	299	
QY	290	DQIIKGLALAMLFGTDSSTCTEWSLNLIINHPEVLKKADELDTVOGDRIILNESDLR	349	
DB	299	SD-I EHLLEDMTAGTDTSSTSTLEWAMELLKNPKTKAKAQAEIDCVIGONGIVEESDIS	357	
QY	350	KLPYLKRIILETRLYPAPFLIIPHVSSEDITIEGFI PRDTIVIINGWGQMQRDPOLWN	409	
DB	358	KLPYLQAVVKETFLRLHTVPVLLI PKRAESDAEILGFVMVLKDTQVLNVNVAIGRDPSVMD	417	
QY	410	ATCFKPERP-----DVEGEKKLVAFNGMRACGEPMANQSVSFTGLLIQCEDW---K	461	
DB	418	PSQFEPERFLGKDMDVRGRDYELTFPGAGRICEGMPLEAMKTVMSLMASLYSFDWKLPK	477	
QY	462	RYSKEKLDWTENNWTLSRLIPLEAM 487		
DB	478	GVLSEDLMDMETFGTLTHKTNPLHAV 503		
 RESULT 9				
ID	C751_PETHY	STANDARD;	PRT;	506 AA.
AC	P48418;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Flavonoid 3',5'-hydroxylase 1 (EC 1.14.-.-) (F3'5'H) (Cytochrome P450 75A1) (CYPLXVALL).			
DN	CYP75A1 OR HFI.			
OS	Petunia hybrida (Petunia).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Asteridae; lamiids; Solanales; Solanaceae; Petunia.			
OX	NCBI_TaxId=4102;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Old Glory Blue; TISSUR=Petal;			
RX	MEDLINE=94050182; PubMed=8232589;			
RA	Holton T.A., Brugliera F., Lester D.R., Tanaka Y., Hyland C.D.,			
RA	Menting J.G.T., Lu C.Y., Farcy E., Stevenson T.W., Cornish E.C.;			
RT	"Cloning and expression of cytochrome P450 genes controlling flower			

RT COLOUR.";
 RL Nature 366:276-279(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Falcon Blue; TISSUE=Petal;
 RA Ohbayashi M., Shimada Y., Nakano R., Kiyokawa S., Kikuchi Y.;
 RT "Molecular cloning of cDNA encoding flavonoid-3',5'-hydroxylase from
 RL Petunia hybrida.";
 RN Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RA Hwang T.S., Chen Y.C., Wu R.Y.;
 RT "Cloning and sequencing of the genomic flavonoid 3',5'-hydroxylase
 RL (F3',5'Hase) gene from Petunia hybrida.";
 RN Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE OF 1-425 FROM N.A.
 RC STRAIN=cv. Blue-star; TISSUE=Flower;
 RA Toguri T., Azuma M., Ohtani T.;
 RT "The cloning and characterization of a cDNA encoding a cytochrome P450
 RL from the flowers of Petunia hybrida.";
 RN Plant Sci. 94:119-126(1993).
 CC -!- FUNCTION: Catalyzes the 3',5'-hydroxylation of naringenin and
 CC eriodictyol to form 5,7,3',4',5'-pentahydroxyflavanone and 3',5'-
 CC hydroxylation of dihydrokaempferol and dihydroquercetin to form
 CC dihydromyricetin.
 CC -!- PATHWAY: Anthocyanin biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial (Potential).
 CC -!- TISSUE SPECIFICITY: FLOWERS.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D22545; CAA80266.1; -;
 DR EMBL; D14588; BAA03438.1; -;
 DR EMBL; AF081575; AAC32274.1; -;
 DR EMBL; X71130; CAA50442.1; -;
 DR PIR; S32110; S32110.
 DR PIR; S38985; S38985.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR PRINTS; PR00385; P450; 1.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Heme; Multigene family.
 FT METAL 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 506 AA; 56548 MW; FCF3F9623BD82B3E CRC64;
 Query Match 29.2%; Score 756; DB 1; Length 506;
 Best Local Similarity 32.1%; Pred. No. 1.8e-36;
 Matches 159; Conservative 114; Mismatches 200; Indels 22; Gaps 5;
 QY 10 LFLVLFLGVKVFQSRKLRNIPPGPPPLIIGNLNLLRQPIHRFFQRMKQYGNVSLWF 69
 DB 13 IFLIAHIIISLTISKTGTHLPGRGWPVIGALPLGAMPHVSLAKWAKYGAIMYLKV 72
 QY 70 GSRLAVVISPAYQECTKIDVALANRLPSLGKYPYNNNTVGSCHGHEWNLRRIT 129
 DB 73 GTCGNMAVASTPDAAKAFKLTLDINFSNPPNAGATHLAYNAQDMVFAHYGPRWLLKLS 132
 QY 130 ALDVLSTORVHSFGIRSDKRLMQLRVLAKNSNEEFARVEISSMENDLTYNIMEMI 189
 DB 133 NLHMGKALENWANVRANELGHMLKSM-----SDMSREGQVVVAEMLTFFAMANNIGQVM 188
 QY 190 SKRFYGESEMKNVVEAREFRETVTETMLEMLGLANKGDHPLFLRWFDFQNVKRLKLSIS 249
 DB 189 LSKRVFVDKG-----VEVNEFKDMVVELMTIAGYFNIGDFPCLAWDLQIEKRMKRLH 243

QY 250 KRYDSTILKILHNENRASNDQRNSMDHLLKLOE-----TOPQYVTDQIIIGLALAMFGCT 305
 DB 244 KKFDFALLTRMFDKHKATTYERKGFDFLDVVMENGNDSEGERLSTNTIKALLINLFTAGT 303
 QY 306 DSTGTGLEWLSNLNLNHPVLKARDELDTQVGDRLNLSLPLKLPYLRKILLETFLRLY 365
 DB 304 DTSSSAIEWALAEWMKNPAILKKAQAEQVIGRRRLLESIPNLPYLRACKETFRKH 363
 QY 366 PPAPILIPHVSSEDTIEGFINPRDTIVINGMGHQRDPQLWNDATCFKPERP----- 418
 DB 364 PSTPLNLPRISEPCIVDGYIPKNTRLSNVIAIGRDPQVWENPLEFNPFRFLSGRNSK 423
 QY 419 -DVEGEKLVAFGMRACPCPEPMQSVSTGLIGLICQFQWKRUSEE-KLDMTENNWI 476
 DB 424 IDPRGNDFLIIPFGARRICAGTRMGIVNVVETILGTLVHSEFDWKLPSEVIELNMEAFGL 483
 QY 477 TSLRLIPLEAMCKAR 491
 DB 484 ALQKAVPLEAMVTPR 498
 RESULT 10
 CP71 PERAE
 ID CP71 PERAE STANDARD; PRT; 471 AA.
 AC P24465;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 71A1 (EC 1.14.-.-) (CYPLXXIAI) (ARP-2).
 GN CYP71A1.
 OS Persea americana (Avocado).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Persea.
 OX NCBI_TaxID=3435;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Hass;
 RX MEDLINE=90251665; PubMed=1692626;
 RA Bozak K.R., Yu H., Sirevag R., Christoffersen R.E.;
 RT "Sequence analysis of ripening-related cytochrome P-450 cDNAs from
 RL avocado fruit.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3904-3908(1990).
 RN [2]
 RP SEQUENCE OF 1-40.
 RC TISSUE=Mesocarp;
 RA O'Keefe D.P., Leto K.J.;
 RT "Cytochrome P-450 from the mesocarp of avocado (Persea americana).";
 RL Plant Physiol. 89:1141-1149(1989).
 CC -!- FUNCTION: INVOLVED IN THE METABOLISM OF COMPOUNDS ASSOCIATED WITH
 CC THE DEVELOPMENT OF FLAVOR IN THE RIPENING FRUIT PROCESS, POSSIBLY
 CC BY ACTING AS TRANS-CINNAMIC ACID 4-HYDROLASE.
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -!- SUBCELLULAR LOCATION: MICROSOMAL; MEMBRANE-BOUND (POTENTIAL).
 CC -!- TISSUE SPECIFICITY: MESOCARP.
 CC -!- DEVELOPMENTAL STAGE: IN RIPENING FRUIT.
 CC -!- INDUCTION: By ethylene.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M32885; AAA22913.1; ALT_TERM.
 DR PIR; A35867; A35867.
 DR HSSP; P14779; IJUZ.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.

DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Heme; Transmembrane; Fruit ripening.
 FT TRANSMEM 7 21 POTENTIAL.
 FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 471 AA; 53074 MW; C7B7ECF506D3EBD CRC64;
 Query Match 29.1%; Score 753; DB 1; Length 471;
 Best Local Similarity 34.9%; Pred. No. 2.4e-36;
 Matches 166; Conservative 98; Mismatches 180; Indels 32; Gaps 8;
 QY 8 AVLEFLVLGVKVFQSRKLR-----NIPGPPPLPIIGNLNLEQPIHRRFFQMSQ 60
 DB 2 AILVSLFLAIALTFLLKLNKREKKNPLPFPNLPPIIGNLHQGNVPHRSLRLANE 61
 QY 61 YGNVSLWFGSLAVVISSPAYQECFTKHDVALANRLPSLSGKYFYNNNTTVGSCSHGE 120
 DB 62 LGPLILLHLGHIPTLIVSTAEIAEILKTHLIFASRPSTTAARRIFYDCTDVAFPYGE 121
 QY 121 HWNLRLRITALDVLSTQVHSFGIRSDETKRLMQLVLAKNSEEFARVEISSMFDNL 180
 DB 122 YWQVRKICVLELLSKRVNSYSIRSEEVGLMWERISQSCSTGE----AVNLSELLLLL 177
 QY 181 TYNIMRMISGKPFYGEESMKVVEAREPREVTWLEMLGLANKGDHLPFLRWFD-FQ 239
 DB 178 SSGTITRVAFGKYESEE-ERKN-----KPADLATELTTLMGAFVGDYFPSPAWVDVLT 231
 QY 240 NVEKRLKSIKRYDSTILNKILHE-----NRASNDRONSMIDHLLKLOETOP--QYTTD 290
 DB 232 GMDARLKRNGHGLDAFVDHVIDDHLLSRKANGSDGVEQKLDVLLHLQKDSGLGVHLNR 291
 QY 291 QIITGLALAMLFQGTDSSTGTLKSLNLLNHPVLEKARDELDTQVQDRLLNESDLPK 350
 DB 292 NNLKAVILDFMFGSGTDTTAVTLEWMAELIKHPDVMKEAQOQEVRRVVGKAKVEEEDLHQ 351
 QY 351 LPYLKILITLALYPPAPILIPHVSEDITIGENIPRDTIVTINGWGMORPQLWMDA 410
 DB 352 LHLVLLIKETLRLHPVAPILVPRESTRDVVIRGYHIPAKTRVFINAWAIGRDPKSWENA 411
 QY 411 TCFKPERF-----DVEGEKKLVAFGMGRACPGCEPMAMQSVFTLGLLIQCQDWK 461
 DB 412 EEFLEPERFVNNSVDFKQDQFOLIPFGAGRGRCGFIAGISSVEISLANLLYWFNWE 467
 RESULT 11
 F3PH PETH STANDARD; PRT; 512 AA.
 AC Q9SB09;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Flavonoid 3'-monooxygenase (EC 1.14.13.21) (Flavonoid 3'-hydroxylase)
 DE (Cytochrome P450 75B2).
 GN CYP75B2 OR HT1.
 OS Petunia hybrida (Petunia).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamiales; Solanales; Solanaceae; Petunia.
 OC NCBI_TaxID=4102;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Old Glory Red; TISSUE=Corolla;
 RX MEDLINE=99435883; PubMed=10504566;
 RA Brugliera F., Barri-Rewell G., Holton T.A., Mason J.G.;
 RT "Isolation and characterization of a flavonoid 3'-hydroxylase cDNA
 RT clone corresponding to the Hc1 locus of Petunia hybrida";
 RL Plant J. 19:441-451 (1999).
 RC -!- FUNCTION: Catalyzes the 3'-hydroxylation of the flavonoid B-ring
 CC to the 3',4'-hydroxylated state. Convert naringenin to eriodictyol
 CC and dihydrokaempferol to dihydroquercetin.
 CC CATALYTIC ACTIVITY: A flavonoid + NADPH + O(2) = 3'-
 CC hydroxyflavonoid + NADP(+) + H(2)O.
 CC -!- PATHWAY: Flavonoid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).

RESULT 12

CC -!- TISSUE SPECIFICITY: High expression in petals and ovaries and to a
 CC lower extent in sepals, pedicels, anthers and stems. Not detected
 CC in leaves, style or roots.
 CC -!- DEVELOPMENTAL STAGE: High expression early in flower development
 CC and then declines as the corolla becomes fully pigmented and
 CC opened.
 CC -!- MISCELLANEOUS: May act as a membrane anchor for localization of
 CC other, soluble, flavonoid enzymes to the endoplasmic reticulum.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF155332; AAD56282.1; --
 CC InterPro; IPR001128; Cytochrome_P450.
 CC Pfam; PF00067; P450; 1.
 CC PRINTS; PR00385; P450.
 CC DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Flavonoid biosynthesis; Oxidoreductase; Monooxygenase;
 KW Endoplasmic reticulum; Heme; NADP; Signal-anchor; Transmembrane.
 FT TRANSMEM 1 21 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT DOMAIN 22 512
 FT CYTOPLASMIC (POTENTIAL).
 FT METAL 447 447
 FT IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 512 AA; 56936 MW; 1843AC1BEB32BAE2 CRC64;
 Query Match 28.8%; Score 746.5; DB 1; Length 512;
 Best Local Similarity 33.3%; Pred. No. 6.3e-36;
 Matches 168; Conservative 104; Mismatches 203; Indels 29; Gaps 9;
 QY 1 LLVVSVYAVL--FLVFLGVKVFQSRKLRNIPGPPPLPIIGNLNLEQPIHRRFFQMS 58
 DB 3 ILSLLIYVIFESFLQFI-LRSFFRKRYPLPLPFGPKPWPILGNLVLHLPKPHQSTAAMA 61
 QY 59 KOYGNVSLWFGSLAVVISSPAYQECFTKHDVALANRLPSLSGKYFYNNNTTVGSCSH 118
 DB 62 QTYGPLYMLKMGFVDVWVAASAAVAAQFLKTHDANFSSRPSPNSGAEHMAYNYQDLVFAPY 121
 QY 119 GEHWNLRLRITALDVLSTQVHSFGIRSDETKRLMQLVLAKNSEEFARVEISSMFDNL 178
 DB 122 GPRWMLRKICSVHLFSTKALDDFRHVRQDEKTLTTRALASAGQK-----PVKUGQLLN 175
 QY 179 DLTYNIMRMISGKPFYGEESMKVVE-FAREFRETVTMELMGLANKGDHLPFLRWFD 237
 DB 176 VCTTVALARKVMIGKRVFADGS--GDVDPQAAEFKSNVEMVAVGVNFGDIPQLNWLD 233
 QY 238 FQNVVKRLKSIKRYDSTILNKILHENRAS--NDRONSMIDHLLKLO---ETQPYTYDQI 292
 DB 234 IQGVAAKMKKLHARFADFALTILEEHKGIKFGEMKDLLSTLSLKNDDANDNGKLTDTYE 293
 QY 293 IKGALAMLFQGTDSSTGTLKSLNLLNHPVLEKARDELDTQVQDRLLNESDLPKLP 352
 DB 294 IKALLNLNLFVAGTDTSSSTVEWAIARLNPKILAAQAOQIEDKVVGRDLVGLDGLAQLT 353
 QY 353 YLRKIITLRLYPPAPILIPHVSEDITIGENIPRDTIVTINGWGMORPQLWMDATC 412
 DB 354 YLEAVKETFRLLHPSTPLSLPRIASESCINGYFPKGSTILLNVAIARPNWADPPL 413
 QY 413 FKPERF-----DVEGEKKLVAFGMGRACPGCEPMAMQSVFTLGLLIQCQDWKRV 463
 DB 414 FRPERLPGGEKPKVDVGRGDFEVIFFGAGRRICAGMNLGIRVMQLMIATLIHAFNWDLV 473
 QY 464 S---BEKLDMTENNITLSRLIPL 484
 DB 474 SGQLPEMLNBBAYGLTLQRAADPL 497


```
DR InterPro; IPR001128; Cytochrome_p450.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Heme; Transmembrane.
FT NON_TER 1 15
FT TRANSMEM <1 15 POTENTIAL.
FT METAL 450 450 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 511 AA; 57954 MW; EFEC238BD8112432 CRC64;

Query Match      28.2%; Score 731.5; DB 1; Length 511;
Best Local Similarity 32.5%; Pred. No. 4.5e-35;
Matches 164; Conservative 104; Mismatches 203; Indels 33; Gaps 10;

Qy 10 LFLVLFLGVKVFVQSRKLRI-----PPGPPPLPIIIGNLLEQPIHRRFFORMSKOYGN 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 ILIALLCTLPFLPKKWRRSYSGKTPPPSPKLPVLGNLHQLGTTPPHRSLSLSRRYGP 60

Qy 64 VVSLWFGSRLAVVSSITAYQECCTKHDVALANRLPSLS-GKYIFYNNTTVGSCSHGEHW 122
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 VMQLHFGSPVLVAVSSPEAREIMKNQDLNFSNR-PNLSIPRLLYDNHDVAFAPYGEYW 119

Qy 123 RNLFRITALLVLTQVHSPSGIRSDETKRLMORLVLAKNSEEEFAARVEISSMENDLTY 182
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 RQIRSIQVQLLSNKRVSQFRVREETSIMVEKIMQLKTTPT--AAVNLDTLLTCLTN 177

Qy 183 NNIMRMISGRKFYGESEEMKNVEEAREFRETVTTEMLMGLANKGDHLPFLRWF-DFQNV 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 178 DVFCRIALGKK-YGTTTGDEYHVRSLKKNLAEFYVLLGISPLWEYIPWLEWTRFDGV 236

Qy 242 EKRLKSISKRYDSILNKLILHENRASNDRONS-----MIDHLKLQ---ETQPOYY 288
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 237 DRRIEVSRTDFELGKVIQEHVRVDRKREDTIVVGDVTVGLDFVDLLQLQFQRENERSSPV 296

Qy 289 TDQIIKGLALAMFGGTDSTGTLEWSLNLAHPEVLKKAARDELDTQVGODRLNEDSL 348
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 297 DDLTIKAVILDMFLAGDTITVTALEWALSELIKNPRAWKILQKEVRGVAGSKGEIESDL 356

Qy 349 PKLPYLRKIILLETLLRYPAPILIPHVSSDITIEGFNIPRDTIIVINGMQMDPOLWN 408
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 357 EKMPYLKAVMKESLRLLHAPVLLVPRESTRTDKVLGYDVASGTRVLINCWAIGRDSVWE 416

Qy 409 DATCFKPERF-----DVEGEKKLVAFGMGRACPGEPMAQSVSFTLGLLIQCDFWK-- 461
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 417 ESETFLPERFLETSIDYRGMHFELIPFGSGRRGCGGATFAAIDELALATLVHKKDFKLP 476

Qy 462 -RVSEKLDMTENNWITLSRLIPL 484
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 477 NGVRVEDLDMSESGSGFTIHKFPL 500
```

Search completed: October 20, 2003, 20:32:09
Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 20:28:34 ; Search time 102 Seconds
(without alignments)
1262.434 Million cell updates/sec

Title: US-09-868-546A-2
Perfect score: 2592
Sequence: 1 LLVWSYAVLVFLVFLGVKFP.....RLIPLEAMCKARPLATKIGI 499

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTRMBL23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1671.5	64.5	499	Q9ZWF2	Q9zwf2 glycyrrhiza
2	1646.5	63.5	498	Q9ZRW6	Q9zrw6 cicer ariet
3	1646.5	63.5	498	Q9XFX0	Q9xfx0 cicer ariet
4	1581	61.0	499	Q9MBE4	Q9mbe4 lotus japon
5	1350	52.1	499	Q9XFX1	Q9xfx1 cicer ariet
6	1287.5	49.7	500	Q23154	Q23154 arabidopsis
7	1285	49.6	499	Q9SZT6	Q9szt6 arabidopsis
8	1276	49.2	509	Q9LHA1	Q9lha1 arabidopsis
9	1241	47.9	497	Q9SZT7	Q9szt7 arabidopsis
10	1229	47.4	492	Q23155	Q23155 arabidopsis
11	1227	47.3	494	Q8LCN8	Q8lcn8 arabidopsis
12	1219.5	47.0	495	Q23156	Q23156 arabidopsis
13	1212	46.8	543	Q22188	Q22188 arabidopsis
14	1211	46.7	501	Q8LA85	Q8la85 arabidopsis
15	1205	46.5	515	Q22185	Q22185 arabidopsis
16	1199.5	46.3	481	Q9C714	Q9c714 arabidopsis

17	1189.5	45.9	491	10	Q9LVD6	Q9lvd6 arabidopsis
18	1188.5	45.9	500	10	Q8LBE7	Q8lbe7 arabidopsis
19	1171.5	45.2	500	10	Q8S9L8	Q8s9l8 arabidopsis
20	1168.5	45.1	518	10	Q9SW67	Q9sw67 arabidopsis
21	1144.5	44.2	484	10	Q8H137	Q8h137 arabidopsis
22	1132	43.7	520	10	Q6S815	Q6s815 helianthus
23	1125	43.4	501	10	Q9SZU1	Q9szu1 arabidopsis
24	1078	41.6	496	10	Q9FN20	Q9fn20 arabidopsis
25	1078	41.6	507	10	Q8GZ20	Q8gz20 arabidopsis
26	1023	39.5	457	10	Q8VVA6	Q8vva6 arabidopsis
27	1019	39.3	732	10	Q94HA5	Q94ha5 oryza sativ
28	1012	39.0	518	10	Q94HA3	Q94ha3 oryza sativ
29	1000	38.6	517	10	Q9ATV6	Q9atv6 lolium rigi
30	999.5	38.6	527	10	Q94HA4	Q94ha4 oryza sativ
31	995.5	38.4	512	10	Q94HA6	Q94ha6 oryza sativ
32	993	38.3	517	10	Q9ATV5	Q9atv5 lolium rigi
33	983	37.9	517	10	Q9ATV4	Q9atv4 lolium rigi
34	979	37.8	414	10	Q940P7	Q940p7 arabidopsis
35	903.5	34.9	561	10	Q9LXB2	Q9lxb2 arabidopsis
36	888.5	34.3	500	10	Q9LXB3	Q9lxb3 arabidopsis
37	840	32.4	372	10	Q81653	Q81653 hemerocalli
38	828.5	32.0	513	10	Q9FPN2	Q9fpn2 matthiola i
39	824	31.8	540	10	Q9FVK6	Q9fvk6 pisum sativ
40	810.5	31.3	509	10	Q94FM3	Q94fm3 nicotiana t
41	808.5	31.2	523	10	Q49394	Q49394 arabidopsis
42	801	30.9	524	10	Q9SZ46	Q9sz46 arabidopsis
43	793.5	30.6	526	10	Q8S845	Q8s845 oryza sativ
44	783	30.2	338	10	Q9SZU0	Q9szu0 arabidopsis
45	782.5	30.2	240	10	Q9SML2	Q9sml2 cicer ariet

ALIGNMENTS

RESULT 1

Q9ZWF2 ID Q9ZWF2 PRELIMINARY; PRT; 499 AA.

AC Q9ZWF2; (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-MAY-2003 (TREMBLrel. 23, Last annotation update)

DE Cytochrome P450.

GN CYP GE-31.

OS Glycyrrhiza echinata (Licorice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosoids I; Fabales; Fabaceae; Papilionoideae; Galegeae; Glycyrrhiza.

OX NCBI_TaxID=46348;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20008263; PubMed=10540749;

RA Nakamura K., Akashi T., Aoki T., Kawaguchi K., Ayabe S.;

RT "Induction of isoflavonoid and retrochalcone branches of the flavonoid

RT pathway in cultured glycyrrhiza echinata cells treated with yeast

RT extract.";

RL Biosci. Biotechnol. Biochem. 63:1618-1620(1999).

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

DR EMBL; AB022732; BAA74465.1; -

DR InterPro; IPR001128; Cytochrome_P450.

DR Pfam; PF00067; P450; 1.

DR PRINTS; PR00385; P450.

DR PROSITE; PS00086; CYTOCHROME_P450; 1.

KW Heme; Monooxygenase; Oxidoreductase.

SQ SEQUENCE 499 AA; 57227 MW; 0F65DA06F88D3D64 CRC64;

Query Match

Best Local Similarity 64.5%; Score 1671.5; DB 10; Length 499;

Matches 304; Conservative 99; Mismatches 91; Indels 3; Gaps 1;

QY 1 LLVWSYAVLVFLVFLGVKFPQSRKRNIPGPPPLPIIGNLNLEQPIHFFORMSKO 60

3 ILLSLSYSVFLALPIFIVIRARKFKNLPGPSPLPIGNLHLKRLHRTFKLSEK 62

QY 61 YGNVSLWFGSLRLAVISSPTAYQECFTKNDVLANRPLSLGKYIFYNNTTGVSCSHGE 120
 Db 63 YGHVSLWFGSLRLVWSSASEFQCCFTKNDVLANRPLSLGKYIFYNNTTGLSTSYGE 122
 QY 121 HWRNLRRITADVLSQVHSGIRSDGTRKLMQRLVLANRPLSLGKYIFYNNTTGLSTSYGE 180
 Db 123 HWRNLRRITADVLSQVHSGIRSDGTRKLMQRLVLANRPLSLGKYIFYNNTTGLSTSYGE 179
 QY 181 TYNMIRMTISGRFYGESEEMKNVEAREFRETTVTEMLMGLANKGDHLPFLRWDFQ 240
 Db 180 TFNNIMRMISGRFYGESEEMKNVEAREFRETTVTEMLMGLANKGDHLPFLRWDFQ 239
 QY 241 VEKRLKISKRYDYLKILHNRASNDQRNSMDHLLKLOETQPOYYTDOIIKGLALAM 300
 Db 240 LEKRLKISKRYDYLKILHNRASNDQRNSMDHLLKLOETQPOYYTDOIIKGLALAM 299
 QY 301 LFGTDSSTGTLEWSNLNHNPEVLKARDELDTQVQDRLNLSDELPLKYLKIL 360
 Db 300 LLAGTDSSTGTLEWSNLNHNPEVLKARDELDTQVQDRLNLSDELPLKYLKIL 359
 QY 361 TRLYPAPILPHVSSSEDIIEGNIIPRDTIVINGMQRDPOLWMDATCFKPERFDV 420
 Db 360 TRLYTPAPILPHVSSSEDIIEGNIIPRDTIVINGMQRDPOLWMDATCFKPERFDV 419
 QY 421 EGEEKLVAFGMGRACGEPGMAMQSVFTGLLIQCFDVKVSEKLDNTENNWTLSR 480
 Db 420 KGEIEKLVAFGMGRACGEPGMAMQSVFTGLLIQCFDVKVSEKLDNTENNWTLSR 479
 QY 481 LIPLEAMCKARPLATKI 497
 Db 480 LVPLKAMCKSRPVNKV 496

RESULT 2

Q9ZRW6 PRELIMINARY; PRT; 498 AA.
 AC Q9ZRW6;
 DT 01-MAY-1999 (TRENBLrel. 10, Created)
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Cytochrome P450 (EC 1.14.14.1).
 GN CYP81E3.
 OS Cicer arietinum (Chickpea) (Garbanzo).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
 OX NCBI_TaxID=3827;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. ILC3279;
 RA Overkamp S., Barz W.;
 RT "Cloning of two Cicer arietinum L. cDNA's encoding Cytochrome P450s
 RT highly homologous to Isoflavone 2'-Hydroxylase from Licorice.";
 RL Plant Physiol. 120:935-935(1999).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AJ012581; CAA10067.1; --
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 498 AA; 57397 MW; 04BDC8E8EBF061D1 CRC64;

Query Match 63.5%; Score 1646.5; DB 10; Length 498;
 Best Local Similarity 59.4%; Pred. No. 8.5e-108;
 Matches 295; Conservative 101; Mismatches 98; Indels 3; Gaps 1;
 QY 1 LLVVSVAVLVLFLGVKVFQSKRLNIPPGPPPLPIGNLNLLEQPIHRRFFQMSKQ 60
 Db 3 ILSLCSLFLYLSFFIRLLFQSKRKNLPPGPPPLPIGNLNLLEQPIHRRFFQMSK 62
 QY 61 YGNVSLWFGSLRLAVISSPTAYQECFTKNDVLANRPLSLGKYIFYNNTTGVSCSHGE 120

Db 63 YGDIIISLWFGSLRLVWVSSLSSEFQCCFTKNDVLANRPLSLGKYIFYNNTTGLSTSYSE 122
 QY 121 HWRNLRRITADVLSQVHSGIRSDGTRKLMQRLVLANRPLSLGKYIFYNNTTGLSTSYSE 180
 Db 123 HWRNLRRITADVLSQVHSGIRSDGTRKLMQRLVLANRPLSLGKYIFYNNTTGLSTSYSE 179
 QY 181 TYNMIRMTISGRFYGESEEMKNVEAREFRETTVTEMLMGLANKGDHLPFLRWDFQ 240
 Db 180 TFNNIMRMISGRFYGESEEMKNVEAREFRETTVTEMLMGLANKGDHLPFLRWDFQ 239
 QY 241 VEKRLKISKRYDYLKILHNRASNDQRNSMDHLLKLOETQPOYYTDOIIKGLALAM 300
 Db 240 LEKRVANIADKDAFLRGLQQRKKERTNTMIDHLLNLQSQPEYYTDOIIKGLALAM 299
 QY 301 LFGTDSSTGTLEWSNLNHNPEVLKARDELDTQVQDRLNLSDELPLKYLKIL 360
 Db 300 LLAGTDSSTGTLEWSNLNHNPEVLKARDELDTQVQDRLNLSDELPLKYLKIL 359
 QY 361 TRLYPAPILPHVSSSEDIIEGNIIPRDTIVINGMQRDPOLWMDATCFKPERFDV 420
 Db 360 TRLYTPAPILPHVSSSEDIIEGNIIPRDTIVINGMQRDPOLWMDATCFKPERFDV 419
 QY 421 EGEEKLVAFGMGRACGEPGMAMQSVFTGLLIQCFDVKVSEKLDNTENNWTLSR 480
 Db 420 KGEIEKLVAFGMGRACGEPGMAMQSVFTGLLIQCFDVKVSEKLDNTENNWTLSR 479
 QY 481 LIPLEAMCKARPLATKI 497
 Db 480 LVPLKAMCKSRPVNKV 496

RESULT 3

Q9XFX0 PRELIMINARY; PRT; 498 AA.
 AC Q9XFX0;
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Cytochrome P450 monooxygenase (EC 1.14.14.1).
 GN CYP81E3V2.
 OS Cicer arietinum (Chickpea) (Garbanzo).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
 OX NCBI_TaxID=3827;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. ILC 3279;
 RA Overkamp S., Barz W.;
 RT "Cloning of two Cicer arietinum L. cDNA's encoding Cytochrome P450s
 RT highly homologous to Isoflavone 2'-Hydroxylase from Licorice.";
 RL Plant Physiol. 120:935-935(1999).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AJ238439; CAB41490.1; --
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 498 AA; 57524 MW; F26675C98F771BD8 CRC64;

Query Match 63.5%; Score 1646.5; DB 10; Length 498;
 Best Local Similarity 60.4%; Pred. No. 8.5e-108;
 Matches 297; Conservative 96; Mismatches 96; Indels 3; Gaps 1;
 QY 6 SYAVLVLVLFLGVKVFQSKRLNIPPGPPPLPIGNLNLLEQPIHRRFFQMSKQGNV 65
 Db 8 AVSLFLYLSFFIRLLFQSKRKNLPPGPPPLPIGNLNLLEQPIHRRFFQMSKTYGDI 67
 QY 66 SLWFGSLRLAVISSPTAYQECFTKNDVLANRPLSLGKYIFYNNTTGVSCSHGEHWRNL 125
 Db 68 SLWFGSLRLVWVSSLSSEFQCCFTKNDVLANRPLSLGKYIFYNNTTGLSTSYSEHWRNL 127

QY 126 RRITADVLTQVHVSFGIRSDTKRMQRLVLAKNSNEEFARVEISSMENDLTNNI 185
 DB 128 RRITADVLTQVHVSFGIRSDTKRMQRLVLAKNSNEEFARVEISSMENDLTNNI 184
 QY 186 MRISGRFYGESESEKMNVEAEFEFRETVEMLMGLANKGHDLPFLWDFQNVKRLK 245
 DB 185 MRISGRFYGESESEKMNVEAEFEFRETVEMLMGLANKGHDLPFLWDFQNVKRLK 244
 QY 246 KSISKRYDSILNKILHNENRASNDRONSMIDHLLKLOETQOYYTDOIIKGLALAMFGGT 305
 DB 245 KNIADKTDAFLGLQEQORNKERTNTMIDHLLNQESQPEYYTDOIIKGLALAMLLACT 304
 QY 306 DSSTGTLEWSLNLNHNPEVLKARDELDTQVQDRLLNESDLPKPYRKILLETIRLY 365
 DB 305 DSSAVTLEWSLNLNHNPEVLKARDELDTQVQDRLLNESDLPKPYRKILLETIRLY 364
 QY 366 PPAPILIPHVSSEDTIEGNIPTDTIIVINGMQRDPOLWMDATCFKPERFVGESEK 425
 DB 365 TPAPILIPHVSSEDTIEGNIPTDTIIVINGMQRDPOLWMDATCFKPERFVGESEK 424
 QY 426 KLVAFGMRGRACGPEPMAMQSVFTGLGLIQCFDWKRVSEBKLDNTENNWITLSRLIPLE 485
 DB 425 KLVAFGMRGRACGPEPMAMQSVFTGLGLIQCFDWKRVSEBKLDNTENNWITLSRLIPLE 484
 QY 486 AMCKARPLATKI 497
 DB 485 AMCKTRPVVVKI 496

RESULT 4
 Q9MBE4 PRELIMINARY; PRT; 499 AA.
 ID Q9MBE4
 AC Q9MBE4
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Cytochrome P450.
 OS Lotus japonicus.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
 OX NCBI_TaxID=34305;
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Root;
 RC PubMed=11164575;
 RX Shimada N., Akashi T., Aoki T., Ayabe S.;
 RA "Induction of isoflavonoid pathway in the model legume Lotus
 RT japonicus: molecular characterization of enzymes involved in
 RL phytoalexin biosynthesis."
 RL Plant Sci. 160:37-47(2000).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AB025016; BAA93634.1; -.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 499 AA; 57357 MW; 59FDIDBD6DA3353D CRC64;

Query Match 61.0%; Score 1581; DB 10; Length 499;
 Best Local Similarity 59.8%; Pred. No. 3.5e-103;
 Matches 294; Conservative 90; Mismatches 104; Indels 4; Gaps 2;
 QY 7 YAVLFLVFLGVFVFSQSKRLNIPGPPPLPIIGNLLEQPIHRRFFQMSKQYGNVVS 66
 DB 9 YSFYVAFALIKFLGSRKFLNPPGTSPLPIIGNLHUKRPLHRRFFQMSKQYGNVVS 69
 QY 67 LWFSGRLAVVISPTAYQECFTKHDVALANRLPSLSKGIYFYNNNTVGSCHGEHWRNLR 126
 DB 69 LWFENRLVVVVSSFADVQECFTKNDVVLNANRPLSGKIYFYNTTLGTSYGEHWRNLR 128
 QY 127 RITALDVLSTQVHVSFGIRSDTKRMQRLVLAKNSNEEFARVEISSMENDLTNNI 186

DB 129 RITSLDVLSNHRINSFSPIRDTETRLIRKLA---EDSAKNFSEVELTSRFFDMTFNNIM 185
 QY 187 RMISGRFYGESESEKMNVEAEFEFRETVEMLMGLANKGHDLPFLWDFQNVKRLK 246
 DB 186 RMISGRFYGESESEKMNVEAEFEFRETVEMLMGLANKGHDLPFLWDFQNVKRLK 245
 QY 247 KSISKRYDSILNKILHNENRASNDRONSMIDHLLKLOETQOYYTDOIIKGLALAMFGGT 305
 DB 246 GISSKTDRLGLLQEQORNKERTNTMIDHLLNQESQPEYYTDOIIKGLALAMLLACT 305
 QY 306 DSSTGTLEWSLNLNHNPEVLKARDELDTQVQDRLLNESDLPKPYRKILLETIRLY 365
 DB 306 DSSAVTLEWSLNLNHNPEVLKARDELDTQVQDRLLNESDLPKPYRKILLETIRLY 365
 QY 366 PPAPILIPHVSSEDTIEGNIPTDTIIVINGMQRDPOLWMDATCFKPERFVGESEK 425
 DB 365 TPAPILIPHVSSEDTIEGNIPTDTIIVINGMQRDPOLWMDATCFKPERFVGESEK 425
 QY 426 KLVAFGMRGRACGPEPMAMQSVFTGLGLIQCFDWKRVSEBKLDNTENNWITLSRLIPLE 485
 DB 426 KLIPFGLGRACGPEPMAMQSVFTGLGLIQCFDWKRVSEBKLDNTENNWITLSRLIPLE 485
 QY 486 AMCKARPLATKI 497
 DB 486 AMCKSRPVVNNV 497

RESULT 5
 Q9XFX1 PRELIMINARY; PRT; 499 AA.
 ID Q9XFX1
 AC Q9XFX1
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Cytochrome P450 (EC 1.14.14.1).
 GN CYP81E2.
 OS Cicer arietinum (Chickpea) (Garbanzo).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Ciceraceae; Cicer.
 OX NCBI_TaxID=3827;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=cv. ILC 3279;
 RC PubMed=10773344;
 RX Overkamp S., Hein F., Barz W.;
 RA "Cloning and characterization of eight cytochrome P450 cDNAs from
 RT chickpea (Cicer arietinum L.) cell suspension cultures."
 RL Plant Sci. 155:101-108(2000).
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AJ239051; CAB43505.1; -.
 DR InterPro; IPR001064; Crystallin.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 499 AA; 56656 MW; E61B3A1D49E5BF55 CRC64;

Query Match 52.1%; Score 1350; DB 10; Length 499;
 Best Local Similarity 50.6%; Pred. No. 6.6e-87;
 Matches 249; Conservative 113; Mismatches 124; Indels 6; Gaps 5;
 QY 9 VLFVFLVFLGVFVFSQSKRLNIPGPPPLPIIGNLLEQPIHRRFFQMSKQYGNVVS 67
 DB 9 IIFSLFLITTKLFLSKRKFCKLPPCPPSPISPIIGNLHUKRPLHRRFFQMSKQYGNVVS 68
 QY 68 WFGSRLAVVISPTAYQECFTKHDVALANRLPSLSKGIYFYNNNTVGSCHGEHWRNLR 127
 DB 69 KFGSQVWVSSASLVEECFTKNDIIFANRHSATKILGYNNNTVGSCHGEHWRNLR 128

```

QY 128 ITADVLTSTORVHSFGIRSDTKRLMORLVLAKNSENEEFARVEISSMFNDITYNNIMR 187
DB 129 ISSIEILSTHRLNSFSEIRKDETRWLGK--LAKSH-KDFTKVELRSLFGLTFTIMR 185
QY 188 MISGRPYGESEKONVEAREFRETVTEMLMGLANKGDHLPFLRWDFQNVKRLKS 247
DB 186 MVGCKREYGESDGTNAEAKFRDMNNEQEFGLGSLNGLDFVPLFKLDFSGGKKLXK 245
QY 248 ISKRYDSILNKLHNENASNDRO-NSMIDHLLKLOETOPOYYTDQIIKGLALAMLFGGTD 306
DB 246 VGEKWDALFOGLVDEHRRNKKKLNLTWIDHLLSLOESOPYYSQIIKGLIMALIIVAGTE 305
QY 307 SSTGTLEWSNLNHNPEVLKKADELDTQVGDRLNENSDLPKLPYLKILLETLRYP 366
DB 306 TSSVTLEWANSNLNHNPELEKAKIELDNHIGOEHLIEEAATKLOYLQNIISLTLRHP 365
QY 367 PAPILIPHVSSEDITIEGFNI PRDTIIVINGWGMORPOLWANDATCFKPERFD-VESEEK 425
DB 366 FVTMLLPHLSHDCITIGYDVRNTMLMNAWAIHRDPNLWADPMSFKPERFENGQDIDG 425
QY 426 KLVAFGMGRACPGEPAMOSVSFTLLGLLIQCFFDWKRVSEKLDMTENNWITLSRIPLIE 485
DB 426 GFTPFGMGRACPGSGIALTLGLTIGSLIQCFFEWKRAIGKEEVDMSGSGTVPKAIPLIE 485
QY 486 AMCKARPLATKI 497
DB 486 AQCKARPIINKI 497

RESULT 6
O23154 ID O23154 PRELIMINARY; PRT; 500 AA.
AC O23154;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Cytochrome P450-like protein.
CYA10.20 OR AT4G37340.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RP Rose M., Terryn N., Vos P., Heijnen L., Mewes H.W., Schueller C.,
RA Chaltatzis N.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;
RA Bevan M., Terryn N., Vos P., Heijnen L., Mewes H.W., Schueller C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; Z99707; CAB16769.1; -.
DR EMBL; AL161591; CAB80400.1; -.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; Cytochrome P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 500 AA; 55788 MW; F33F5D15A78A413F CRC64;

Query Match 49.7%; Score 1287.5; DB 10; Length 500;
Best Local Similarity 49.7%; Pred. No. 1.7e-82;
Matches 247; Conservative 106; Mismatches 129; Indels 15; Gaps 9;

QY 9 VLELVFLGVKVFV---QSRKLNIPGPP-PLPIGNLNLLEQPIHRRFFORMSKQYGN- 63
DB 6 LIFTFLFSLTFTIGRIKRRPNLPFSPSWALPVIGHRLRLKPLHRRVFLSVSESLGDA 65

```

```

QY 64 -VVSWMFSGRLAVVISSPAYQECFTKHDVALANRLPSLSGKYIFYNNTTVGSCSHGEHW 122
DB 66 PIISLRNLGNLVFVSSHSLAEBCFTKNDVLANRNSLASCHISYGCTTVTVTASYGDHW 125
QY 123 RNLRIITADVLTSTORVHSFGIRSDTKRLMORLVLAKNSENEEFARVEISSMFNDITY 182
DB 126 RNLRIIGAVEIFSAHRLNSFSEIRKDETRWLGK--LAKSH-KDFTKVELRSLFGLTFT 182
QY 183 NNIMISGRPYGESEKONVEAREFRETVTEMLMGLANKGDHLPFLRWDFQNVKRLKS 242
DB 183 NNIMISGRPYGESEKONVEAREFRETVTEMLMGLANKGDHLPFLRWDFQNVKRLKS 242
QY 243 KKLKSIKRYDSILNKLHNENASND-RQNSMIDHLLKLOETOPOYYTDQIIKGLALAML 301
DB 239 KRKKIASRLDEFLQGLVDERREGKEKQNTWMDHLLCLQETQPEYYTNDIIKGLMSLI 298
QY 302 FCGTSSGTGLEWSNLNHNPEVLKKADELDTQVGDRLNENSDLPKLPYLKILLET 361
DB 299 LAGTDTSAVTLTEWTLNHNPEVLKKADELDTQVGDRLNENSDLPKLPYLKILLET 358
QY 362 LRLYPAPILIPHVSSEDITIEGFNI PRDTIIVINGWGMORPOLWANDATCFKPERFDVE 421
DB 359 LRLYPAPILIPHVSSEDITIEGFNI PRDTIIVINGWGMORPOLWANDATCFKPERFDVE 418
QY 422 GEEKLVAFGMGRACPGEPAMOSVSFTLLGLLIQCFFDWKRVSEKLDMTENNWITLSR 480
DB 419 GEAKLVAFGMGRACPGEPAMOSVSFTLLGLLIQCFFDWKRVSEKLDMTENNWITLSR 478
QY 481 LIPLAMCKARPLATKI 497
DB 479 AIPLVAMCKARPVVGKI 495

RESULT 7
O23276 ID O23276 PRELIMINARY; PRT; 499 AA.
AC O23276;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cytochrome P450-like protein.
FE617.10 OR AT4G37360.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RP Rose M., Terryn N., Vos P., Heijnen L., Mewes H.W., Schueller C.;
RA Bevan M., Terryn N., Vos P., Heijnen L., Mewes H.W., Schueller C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP Rose M., Terryn N., Vos P., Heijnen L., Mewes H.W., Schueller C.;
RA Bevan M., Terryn N., Vos P., Heijnen L., Mewes H.W., Schueller C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RN SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AL035601; CAB38203.1; -.
DR EMBL; AL161591; CAB80400.1; -.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; Cytochrome P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.

```


DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 495 AA; 55882 MW; 083BEF6772741AF5 CRC64;

Query Match 47.08; Score 1219.5; DB 10; Length 495;
 Best Local Similarity 47.3%; Pred. No. 1e-77;
 Matches 232; Conservative 100; Mismatches 147; Indels 11; Gaps 6;

QY 11 FLVLFLGVKVFV-OSRKLRIIPGPP-PLPIIGNLNLEQPIHFFORMSKQYCN--VVS 66
 DB 11 FLFLTISIKULLTSKRNKPNLPSPAYPLPVIGHLLHKQPVHRTFHSISKSLGNAPIFH 70

QY 67 LWFSGRLAVISSPTAYOEFTKHDVALANRLPSLSKGIYFYNNNTVSSSHGHEWNLRL 126
 DB 71 LRLGNRLVYISSHSIAEECTKNDVVLANRPDIIMAKHVGYNFTNMAAASGDHWNLRL 130

QY 127 RITALDVLSTORVHSFGSIRSDETKRLMORVLAKNSNEEFARVEISSMFDNLTYNNIM 186
 DB 131 RIAAVEIFSSHRISTFSIRKDEIRRLITHL---SRDSLHGFEVEVLSKSLTLNLAFNII 187

QY 187 RMISGKRFYGBSEMKNVAEAREFRTVTEMLEMLGLANKGDHLPFLRWDFQNVKRLK 246
 DB 188 MWAGKRYVGTGE--DNDEAKVRELIABEIMAGAGSGNLADYLPISNW--VTNFENQTK 243

QY 247 SISRYDSILNKILHENRASNDRONSMDHLLKLOETPOQYTDQIITKGLALAMFGTGD 306
 DB 244 ILGNRLDRVLQKLVDEKRAEKEGQTLIDHLLSPQETPEYTDVITIKGILALVAGTD 303

QY 307 SSTGLWLSNLNHPVLKARDELDTQVGDRLNEDSLPKLYRLKILLETLLRYP 366
 DB 304 TSSVTLWMSNLNHPVLKARDELDTQVGDRLNEDSLPKLYRLKILLETLLRYP 363

QY 367 PAPILIPHVSSDITIEGFNIPRTIVIIINGWQMDPQLMNDATCFKPERFDVEGEKK 426
 DB 364 AVPLLLPHFSSDECKVAGYDMPRTLLTNVWAMHRDPLWEEPERFKPERFEKEGEAR 423

QY 427 LVAFGMRRACPGPMAMQSVFTGLLIQCFDWRKVRSEKLDNTENNITSLRILPLEA 486
 DB 424 LMPFGMRRACPGALGKRLVSLALGCLIQSFEWERYGAELVDMDTEGEGITMPKATPLRA 483

QY 487 MCKARPLATK 496
 DB 484 MCKARAIVGK 493

RESULT 13
 O22188 PRELIMINARY; PRT; 543 AA.
 AC O22188;
 DT 01-JAN-1998 (TRENBLrel. 05, Created)
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Putative cytochrome P450.
 GN AT2G23190
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vankken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nierman W.C., Venter J.C.,
 RA Salzberg S.L., Fraser C.M., Nierman W.C., White O., Eisen J.A.,
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana".
 RL Nature 402:761-768 (1999).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RL Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 DR EMBL; AC002391; AAB87112.1; -.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME P450; 1.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 543 AA; 62777 MW; 06332DC090FDB12 CRC64;

Query Match 46.8%; Score 1212; DB 10; Length 543;
 Best Local Similarity 47.1%; Pred. No. 4e-77;
 Matches 235; Conservative 102; Mismatches 146; Indels 16; Gaps 9;

QY 3 VVVSVAVLVLVLGVKVFQSRKLR-NIPGPP-PLPIIGNLNLEQPIHFFORMSKQ 60
 DB 50 LILSLAFLF--FISLKLFGKRHSKFNLPSPAPRPLPFIGHLLHKQPLHRTFLFSQS 106

QY 61 YGN--VVSWFGSRLAVISSPTAYOEFTKHDVALANRLPSLSKGIYFYNNNTVSSCSH 118
 DB 107 LGDAPFSLRLGNHLLTVVSSYSIAEECTKNDIVLANRPKFLGKHIEYNFTTWSAPY 166

QY 119 GEHWRNLRITALDVLSTORVHSFGSIRSDETKRLMORVLAKNSNEEFARVEISSMFN 178
 DB 167 GDHWRNLRITGLTLEIFSSHKLNGFLSVRKDEIRHLLLR--LSKNS--QHGFAKVMRQLFY 223

QY 179 DITYNNIMMISGKRFYGBSEMKNVAEAREFRTVTEMLEMLGLANKGDHLPFLWRF-D 237
 DB 224 DLTIINNLMVAGKRFYGEFTQDEY--ARRVQLIDEIVYRAGVGNAAADYIPILSWITD 281

QY 238 FQNVKRLKASISKRYDSILNKILHENRASNDRONSMDHLLKLOETPOQYTDQIITKGLA 297
 DB 282 F--EKGVELASRVDFLOSLVDERRVHKQKNTWMDHLLSLOETPOQYTDVITLKGI 338

QY 298 LAMLFQGTDSSTGLWLSNLNHPVLKARDELDTQVGDRLNEDSLPKLYRLK 357
 DB 339 IVMILAGTETLAGTLEWAMNLNHPVLKARDELDTQVGDRLNEDSLPKLYRLK 398

QY 358 ILETLLRYPAPILIPHVSSDITIEGFNIPRTIVIIINGWQMDPQLMNDATCFKPER 417
 DB 399 VLETLLRYPVAPTNIPHTSDDCILAGYDVPGRGSMLLVNVWMSHRDPSIWEAPEMFKPER 458

QY 418 EDVGEKEKLVAFGMRRACPGPMAMQSVFTGLLIQCFDWRKVRSEKLDNTENNIT 477
 DB 459 FKNEKLNQKLSLFGFGRRACPGVGLAHLRLMSLALGSMVQCFEWRIGEEYVDTRBPMM 518

QY 478 LSRILPLEAMCKARPLATK 496
 DB 519 MRPATPLLAMCKARPIVHK 537

RESULT 14
 O22188 PRELIMINARY; PRT; 501 AA.
 AC O22188;
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Cytochrome P450 monooxygenase-like protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.,
 RT "Full-length messenger RNA sequences greatly improve genome

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 20:21:13 ; Search time 87 Seconds

(without alignments)
910.397 Million cell updates/sec

Title: US-09-868-546A-2

Perfect score: 2592

Sequence: 1 LUVVSYAVFLVFLGVKFP.....RLIPLEAMCKARPLATKIGI 499

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2592	100.0	499	21	AAV96593 Soybean isoflavone
2	2214	85.4	494	21	AAV96594 Soybean isoflavone
3	1251.5	48.3	498	21	AAV50170 Arabidopsis thalia
4	1241	47.9	497	21	AAV32080 Arabidopsis thalia
5	1241	47.9	497	21	AAV50195 Arabidopsis thalia
6	1241	47.9	497	24	ABP81276 Arabidopsis thalia
7	1229	47.4	492	21	AAV30128 Arabidopsis thalia
8	1229	47.4	514	21	AAV30127 Arabidopsis thalia
9	1227	47.3	494	21	AAV26985 Arabidopsis thalia

10	1216	46.9	501	21	AAV50190 Arabidopsis thalia
11	1216	46.9	516	21	AAV50189 Arabidopsis thalia
12	1211	46.7	501	21	AAV44324 Arabidopsis thalia
13	1211	46.7	516	21	AAV44323 Arabidopsis thalia
14	1196.5	46.2	502	21	AAV42048 Arabidopsis thalia
15	1196.5	46.2	505	21	AAV42047 Arabidopsis thalia
16	1188.5	45.9	500	21	AAV16753 Arabidopsis thalia
17	1188.5	45.9	500	21	AAV42954 Arabidopsis thalia
18	1188.5	45.9	519	21	AAV16752 Arabidopsis thalia
19	1188.5	45.9	519	21	AAV42953 Arabidopsis thalia
20	1125	43.4	501	21	AAV12816 Arabidopsis thalia
21	1125	43.4	501	21	AAV43034 Arabidopsis thalia
22	1125	43.4	509	21	AAV12815 Arabidopsis thalia
23	1125	43.4	509	21	AAV43033 Arabidopsis thalia
24	1123	43.3	505	20	AAV05902 Jerusalem artichok
25	1100.5	42.5	457	21	AAV16754 Arabidopsis thalia
26	1100.5	42.5	457	21	AAV42955 Arabidopsis thalia
27	1099	42.4	499	19	AAV79756 Euphorbia lagascae
28	1080	41.7	429	21	AAV42049 Arabidopsis thalia
29	1050.5	40.5	383	21	AAV50171 Arabidopsis thalia
30	1040	40.1	382	21	AAV32081 Arabidopsis thalia
31	1040	40.1	382	21	AAV50196 Arabidopsis thalia
32	1033	39.9	379	21	AAV26986 Arabidopsis thalia
33	921.5	35.6	514	22	AAU02833 Taxus cuspidata ox
34	919	35.5	326	21	AAV50172 Arabidopsis thalia
35	908.5	35.1	325	21	AAV32082 Arabidopsis thalia
36	908.5	35.1	325	21	AAV50197 Arabidopsis thalia
37	901.5	34.8	322	21	AAV26987 Arabidopsis thalia
38	890.5	34.4	319	21	AAV30129 Arabidopsis thalia
39	861	33.2	347	21	AAV50191 Arabidopsis thalia
40	858	33.1	347	21	AAV44325 Arabidopsis thalia
41	850.5	32.8	513	18	AAV35706 Tobacco cytochrome
42	810.5	31.3	509	24	AAV30011 Arabidopsis thalia
43	808	31.2	347	21	AAV12817 Arabidopsis thalia
44	808	31.2	347	21	AAV43035 Arabidopsis thalia
45	790.5	30.5	508	18	AAV35711 Chrysanthum flavon

ALIGNMENTS

RESULT 1

AAV96593
ID AAV96593 standard; Protein; 499 AA.

XX AC AAV96593;

XX XX 26-SEP-2000 (first entry)

XX DE Soybean isoflavone-2-hydroxylase.

XX KW Soybean; isoflavone-2-hydroxylase; flavonol; biosynthesis; anthocyanin;
KW flower colour; pollen tube; feeding deterrent; UV irradiation.

XX OS Glycine max.

XX XX WO200037652-A2.

XX XX 29-JUN-2000.

XX XX 20-DEC-1999; 99WO-US30337.

XX XX 21-DEC-1998; 98US-0113190.

XX XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX XX Famodu OO, McGonigle B, Odell JT, Fader GM, Falco SC;

XX XX WPI; 2000-442678/38.

XX XX N-PSDB; AAA29326.

XX XX New polynucleotide encoding flavonoid biosynthetic enzymes, useful for
PT producing transgenic plants and immunological screening of cDNA

PT libraries

Claim 10; Page 31-32; 36pp; English.

XX This enzyme, a plant (soybean) isoflavone-2-hydroxylase, was encoded by
 CC cDNA, isolated from clone sisc.pk005.n3. It was determined using the
 CC sequence of an isoflavone-2-hydroxylase encoded by a contig composed of
 CC clones sgic.pk001.g17, sg2c.pk004.h7 and slf1.pk0034.g1. The cDNA
 CC sequences can be used for the recombinant production of the enzyme, to
 CC isolate homologues, to create transgenic plants and to provide probes for
 CC genetically and physically mapping genes and as markers for traits linked
 CC to the genes. The proteins can be used for immunological screening, in
 CC particular to raise antibodies against the enzymes. The enzyme and its
 CC gene are useful to study flavonol biosynthesis in plants and provide
 CC means to enhance or otherwise alter flavonol and anthocyanin
 CC biosynthesis. Flavonoids have diverse functions, such as co-pigments in
 CC flower colour, stimulation of pollen tube growth, pollinator attraction,
 CC and feeding deterrents and protection against UV irradiation in fruits
 CC and seeds.

SQ Sequence 499 AA;

Query Match 100.0%; Score 2592; DB 21; Length 499;

Best Local Similarity 100.0%; Pred. No. 4.8e-226;

Matches 499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LLVVSVAVLFLVLFLGVKVFQSRKLRNIPGPPPLPIIGNLLEQPIHRFFORMSKQ 60
 Db 1 LLVVSVAVLFLVLFLGVKVFQSRKLRNIPGPPPLPIIGNLLEQPIHRFFORMSKQ 60
 Qy 61 YGNVSVLWFGSLAVVSSPTAYQECFTKHDVALANRLPSLGKGYFYNNNTVSGSHGE 120
 Db 61 YGNVSVLWFGSLAVVSSPTAYQECFTKHDVALANRLPSLGKGYFYNNNTVSGSHGE 120
 Qy 121 HWRNLRRTALDVLSTQRVHFSFGIRSDTKRLQRLVLAKNSNEEFARVEISSMNDL 180
 Db 121 HWRNLRRTALDVLSTQRVHFSFGIRSDTKRLQRLVLAKNSNEEFARVEISSMNDL 180
 Qy 181 TYNMIRMISGRKFYGESEEMKNVEAREPRETVMLEMLGKNGDHLPLFRWDFQFN 240
 Db 181 TYNMIRMISGRKFYGESEEMKNVEAREPRETVMLEMLGKNGDHLPLFRWDFQFN 240
 Qy 241 VEKRLKSTSKRYDSTLNKILHNENRASNDQRNSMIDHLLKLOETQPOYYTDOIIKGLALAM 300
 Db 241 VEKRLKSTSKRYDSTLNKILHNENRASNDQRNSMIDHLLKLOETQPOYYTDOIIKGLALAM 300
 Qy 301 LFGCTDSSTGTLEWSLNLNHPVLKKADELDTQVGDRLINESDLPKLPYLKRIILE 360
 Db 301 LFGCTDSSTGTLEWSLNLNHPVLKKADELDTQVGDRLINESDLPKLPYLKRIILE 360
 Qy 361 TLRLYPPAPILIPHVSSEDIITIEGFNIPRDTIVINGMGMORDPOLWMDATCFKPERFDV 420
 Db 361 TLRLYPPAPILIPHVSSEDIITIEGFNIPRDTIVINGMGMORDPOLWMDATCFKPERFDV 420
 Qy 421 EGEKKLVAFNGRRACGEPAMQSVFTGLLIQCPDWKRVSEKLDMTENNWITLSR 480
 Db 421 EGEKKLVAFNGRRACGEPAMQSVFTGLLIQCPDWKRVSEKLDMTENNWITLSR 480
 Qy 481 LIPLEAMCKARPLATKIGI 499
 Db 481 LIPLEAMCKARPLATKIGI 499

RESULT 2

AA96594

ID AA96594 standard; Protein; 494 AA.

XX

AC AA96594;

XX

XX 26-SEP-2000 (first entry)

XX

DE Soybean isoflavone-2-hydroxylase.

XX

KW Soybean; isoflavone-2-hydroxylase; flavonol; biosynthesis; anthocyanin;
 KW flower colour; pollen tube; feeding deterrent; UV irradiation.
 XX Glycine max.
 OS
 XX WO200037652-A2.
 PN
 XX 29-JUN-2000.
 PD
 XX 20-DEC-1999; 99WO-US30337.
 PF
 XX 21-DEC-1998; 98US-0113190.
 PR
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA
 XX Famodu OO, McGonigle B, Odell JT, Fader GM, Falco SC;
 PI
 XX WPI; 2000-442678/38.
 DR
 XX N-PSDB; AAA29327.
 XX
 XX New polynucleotide encoding flavonoid biosynthetic enzymes, useful for
 PT producing transgenic plants and immunological screening of cDNA
 PT libraries
 PT
 XX Claim 10; Page 33-34; 36pp; English.
 PS
 XX
 CC This enzyme, a plant (soybean) isoflavone-2-hydroxylase, was encoded by
 CC cDNA, isolated from clone src3c.pk005.f5. It was determined using the
 CC sequence of an isoflavone-2-hydroxylase encoded by a contig composed of
 CC clones sgic.pk001.g17, sg2c.pk004.h7 and slf1.pk0034.g1. The cDNA
 CC sequences can be used for the recombinant production of the enzyme, to
 CC isolate homologues, to create transgenic plants and to provide probes for
 CC genetically and physically mapping genes and as markers for traits linked
 CC to the genes. The proteins can be used for immunological screening, in
 CC particular to raise antibodies against the enzymes. The enzyme and its
 CC gene are useful to study flavonol biosynthesis in plants and provide
 CC means to enhance or otherwise alter flavonol and anthocyanin
 CC biosynthesis. Flavonoids have diverse functions, such as co-pigments in
 CC flower colour, stimulation of pollen tube growth, pollinator attraction,
 CC and feeding deterrents and protection against UV irradiation in fruits
 CC and seeds.

XX SQ Sequence 494 AA;
 Query Match 85.4%; Score 2214; DB 21; Length 494;
 Best Local Similarity 84.4%; Pred. No. 9.2e-192;
 Matches 417; Conservative 45; Mismatches 28; Indels 4; Gaps 3;
 Qy 5 VSYAVLFLVLFLGVKVFQ-SRKLNIPGPPPLPIIGNLLEQPIHRFFORMSKOYGN 63
 Db 1 LSYLLSLVFFFTLKYLFQSRKVRNLPFGPTPLPIIGNLNLVEQPIHRFFORMSKOYGN 60
 Qy 64 VVSLWFGSLAVVSSPTAYQECFTKHDVALANRLPSLGKGYFYNNNTVSGSHGEHWR 123
 Db 61 IISLWFGSLAVVSSPTAYQECFTKHDVTLANRVSLSGKGYFYDNTTVSGSHGEHWR 120
 Qy 124 NLRRTALDVLSTQRVHFSFGIRSDTKRLQRLVLAKNSNEEFARVEISSMNDLTYN 183
 Db 121 NLRRTSLDVLSTQRVHFSFGIRSDTKRLIHR--LARDSG-KDFARVEVTSKFDLTYN 177
 Qy 184 NIMRMISGRKFYGESEEMKNVEAREPRETVMLEMLGKNGDHLPLFRWDFQFVNEK 243
 Db 178 NIMRMISGRKFYGESEELANVEAREFRDVTNEMQLMGLANKGDHLPLFRWDFQFVNEK 237
 Qy 244 RLKSIKRYDSTLNKILHNENRASNDQRNSMIDHLLKLOETQPOYYTDOIIKGLALAMLF 303
 Db 238 RLKSIKRYDSTLNKILHNENRASNDQRNSMIDHLLKLOETQPOYYTDOIIKGLALAMLF 297
 Qy 304 GTDSTGTLEWSLNLNHPVLKKADELDTQVGDRLINESDLPKLPYLKRIILETLR 363
 Db 298 GTDSTGTLEWSLNLNHPVLKKADELDTQVGDRLINESDLPKLPYLKRIILETLR 357
 Qy 364 LYPPAPILIPHVSSEDIITIEGFNIPRDTIVINGMGMORDPOLWMDATCFKPERFDVEGE 423


```

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152263.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 48.3%; Score 1251.5; DB 21; Length 498;
Best Local Similarity 48.3%; Pred. No. 1.9e-104;
Matches ,239; Conservative 110; Mismatches 133; Indels 13; Gaps 6;

QY 9 VLFVLFVGVKVFQSRKLR--NIPGPP-PLFIIGNLNLLEQPIHRRFORMSKQYGN- 63
DB 6 LIFSILFVVLISLIYLIGLKKRKNLPPSPAMSLPVIGHRLRLKKPIHRTFLSLSSQSLNNA 65
QY 64 -VVSIMFGSRLAVISSPTAYQECFTKHDVALANRLPSLSCKYIFYNNTTVGSCSHGHW 122
DB 66 PIFSRLIGNRLVFNSSHSIAECEFKNDDVVLARNPNFILAHHVAYDTTIIAASYGDHW 125
QY 123 RNLRRITDALDVLSTQVHFSFGIRSDETKRLMORLVLAKNSEEBEFARVEISSMENDITY 182
DB 126 RNLRRIGSVFISNHLNLSFLSIRKDEIRLVFRL---SRNFSQEFVVDKMSMLSDLTF 182
QY 183 NNIMRMISGRFYGESEMKVVEAREPRETVTEMLEMLGLANKGDHLPFLRWFDFQVVE 242
DB 183 NNILRMVAGKRYGQDVE--DDPEAKRVROIADVACAGAGNAVLYLFLVRL--VSDYE 238
QY 243 KRLKSIKRYDSIILNKILHENRASNDRONSIMIDHLLKLQETOPOVYTTQIIKGLALAMLF 302
DB 239 TRVKIAGRLDEFLQGLVDEKREAKEKNTWIDHLLTLTLOESQPDYFTDRIIKGNMLSLIA 298
QY 303 GGTDSTGTLEWSLNLNHPVLKKADELDTQVGDRLNEDSLPKLPYLRKIILETL 362
DB 299 GGTDTTAVTLEWALSLLNPNPEVLNKADEIDRMIGVDRLLSESDIPNLPLYQNIVSETL 358
QY 363 RLYPPAPILIPHVSSEDTITIEGFNIPRDTIIVINGMGQRDPOLWNDATCEKPERFDEG 422
DB 359 RLYPAAPMLLPHVASKCKCKGVGYDMPRGTMILLTNAAWTHRDPLLLWDDPTSFKEPERFEKG 418
QY 423 EEKLVAFGMGRACPGEPMAQSVFTLGLLIQCFDWMKRVSEKLDMTENNWITLSRLI 482
DB 419 EAKLMPFGLGRRACPGSLAQRLVTLSGLSIQCFEWERIGEEVDVTEGPGLTMPKAR 478
QY 483 PLEAMCKARPLATKI 497
DB 479 PLEAMCRARDVFGKI 493

RESULT 4
AAG32080
ID AAG32080 standard; Protein; 497 AA.
XX
AC AAG32080;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38634.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.

```

PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140595.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149923.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.

PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-01334256.
PR 14-MAY-1999; 99US-01334218.
PR 14-MAY-1999; 99US-01334219.
PR 14-MAY-1999; 99US-01334221.
PR 14-MAY-1999; 99US-01334370.
PR 18-MAY-1999; 99US-01334768.
PR 19-MAY-1999; 99US-01334941.
PR 20-MAY-1999; 99US-01335124.
PR 21-MAY-1999; 99US-01335353.
PR 24-MAY-1999; 99US-01335629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139419.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140623.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.

PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	23-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161932.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match 47.4%; Score 1229; DB 21; Length 492;			
Best Local Similarity 47.4%; Pred. No. 2.le-102;			
Matches 237; Conservative 107; Mismatches 138; Indels 18; Gaps 7;			
Qy	2	LVVSVYAVLF-LVLFGLVKVFQSKRLRNIPPGPP-PLPIIGNLNLLLEQPIHRPFQRMKS	59
Db	6	ILILSFLSLFLAIFL----LTRSKRLNPPSPALISLPVIGHLLHLLKPLHRTFLSLK	61
Qy	60	QYGN--VVSLWFGRLAVISSPAYOCFTKHDVALANRLPSLGSKYIFYNNTTVGSCS	117
Db	62	SIGNAPVFLRLGNRLVVISRSIAEBCFTKNDVLANRPKFTISKHLGYNATYLLSAS	121
Qy	118	HGEHRLNRRITALDVLSTQVHSFGSIRSDTKRLMQLRLVLAKNSNEEFARVEISSMF	177
Db	122	YGDHWRNLRRAAEIFSTHRLNSFLYIRKDEIRRLSHL---SRDSLHGFVEVEMKTL	178
Qy	178	NLTNNIMRMISGRFYGESEEMKNVVEAREFREITVTEMLEMLGLANKGDHLPFLRFWD	237
Db	179	TNLASNTIIRMLACKRVGEDN-----DDAKLVKLVSEAVTSAGAGNPIDYLSILRW--	231
Qy	238	FQNVKRLKSISKYDSTLNKLHNRASTRONSMIDHLLKLOETQPYITDQIIKGLA	297
Db	232	VSSYEKRIKILGNRFDFTFLQKLVDEKKAKEGETMIDHLLALQIDPYVYTVIIRGII	291
Qy	298	LAMLFGGTDSGTGLENSLNHPEVLKARDELTOVCQDRLLNESDLPLPYLRKI	357
Db	292	LTLLIAGTDSSTVLEWAMSNLLHPEILKARMEIDEKVLGDLRDLVDESIVNLSYLSQSI	351
Qy	358	ILETLRYPAPILIPHVSSEDITIEGNIPTDITVIINGMQRDPOLMNDATCFKPER	417
Db	352	VLETLMYPAPVLLPLHLSSEDCKVGVDIPSGTMTLNTAMHRDPEVWEDPEIFKPER	411
Qy	418	FVGEKEKLVAFGMRACPGEPNMQSVSFTGLLIQCDFMKRVSEKLDMTENNWT	477
Db	412	FEKEGEAEKLISFGMRACPGAGIAHRLINQALGSLVQCPEWERVGEDFVDMTEDKGAT	471
Qy	478	LSRLIPLKAMCKARPLATKI	497
Db	472	LPKAIPLRAMCKARSIVDKL	491
RESULT 8			
ID	AAG30127 standard; Protein; 514 AA.		
XX	AC	AAG30127;	
XX	AC	AAG30127;	
DT	17-OCT-2000 (first entry)		
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 35962.		
XX	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
XX	Arabidopsis thaliana.		
OS	Arabidopsis thaliana.		

XX	EP1033405-A2.	
PN	06-SEP-2000.	
PD	25-FEB-2000; 2000EP-0301439.	
XX	25-FEB-1999;	99US-0121825.
PR	05-MAR-1999;	99US-0123180.
PR	09-MAR-1999;	99US-0123548.
PR	23-MAR-1999;	99US-0125788.
PR	25-MAR-1999;	99US-0126264.
PR	29-MAR-1999;	99US-0126785.
PR	01-APR-1999;	99US-0127462.
PR	06-APR-1999;	99US-0128234.
PR	08-APR-1999;	99US-0128714.
PR	16-APR-1999;	99US-0129845.
PR	19-APR-1999;	99US-0130077.
PR	21-APR-1999;	99US-0130449.
PR	23-APR-1999;	99US-0130510.
PR	28-APR-1999;	99US-0130891.
PR	30-APR-1999;	99US-0131449.
PR	30-APR-1999;	99US-0132048.
PR	04-MAY-1999;	99US-0132407.
PR	05-MAY-1999;	99US-0132484.
PR	06-MAY-1999;	99US-0132485.
PR	06-MAY-1999;	99US-0132486.
PR	07-MAY-1999;	99US-0132487.
PR	11-MAY-1999;	99US-0132863.
PR	14-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	18-MAY-1999;	99US-0134370.
PR	19-MAY-1999;	99US-0134768.
PR	20-MAY-1999;	99US-0134941.
PR	21-MAY-1999;	99US-0135124.
PR	24-MAY-1999;	99US-0135229.
PR	25-MAY-1999;	99US-0135629.
PR	27-MAY-1999;	99US-0136021.
PR	28-MAY-1999;	99US-0136392.
PR	01-JUN-1999;	99US-0136782.
PR	03-JUN-1999;	99US-0137222.
PR	04-JUN-1999;	99US-0137528.
PR	07-JUN-1999;	99US-0137502.
PR	08-JUN-1999;	99US-0137724.
PR	10-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	14-JUN-1999;	99US-0138847.
PR	16-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	21-JUN-1999;	99US-0139763.
PR	22-JUN-1999;	99US-0139817.
PR	23-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.

PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

```

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161362.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 46.9%; Score 1216; DB 21; Length 501;
Best Local Similarity 47.1%; Pred. No. 3.2e-101;
Matches 235; Conservative 117; Mismatches 127; Indels 20; Gaps 11;

QY 9 VLFLVLF-LGVKVFQSKLR-NIDPGPP-PLPIIGNLNLLEQPHREFFORMSKOYGVV 65
DB 7 VLPLALFLAYLFFTSKTRFLNPPSPYSPILGHHNLLKPPVHRLFHRUSKTHGP 66

QY 66 SLWFGSRRLAVISSPAYOECT-KHDVALANRLPSLSGKIYFYNNITVSGCSHGHRN 124
DB 67 SLQFGSRRAVIVSSSLAQCTQGTNDILSNRPCFLTAKYVAYNVTGTAPYGDHWN 126

QY 125 LRRITADVLSQVHSGIRSDETKRLMQRLVLAKNSNEEFARVEISSMFDLTNN 184
DB 127 LRRICSLLELSSNRLTNFLHRRKDETHRLMTR--LSRDVYKKE----IELEPLLSDLTFN 180

QY 185 IWRMLSGKFYGEESMKVGEAREPRETVTEMLGLMGLANKGDHLPFLRWDFONVEKR 244
DB 181 IWRMTGKRYGD--EVHNEEENAVFKLVADINDCSGARHPGDYLPFMKFG-GSFEKK 237

QY 245 LKSIKRYDSILNKLILHENRASNDRQNSMIDHLLKLQETQPOVYTDQITKGLALAMFG 304
DB 238 VKALAEAMDELLORLECKRKD-GNTVWNHLLSQONEPEYTDVTIKGLMGLMWIAG 296

QY 305 TDSSTGLSWSNLNHPVULKARDELDTQVQDRLLNESDLPKLPYLRKILLETURL 364
DB 297 TDTSVTLWAMSSLNHPLEALEKAKLEIDEKIGQERLIDEPDIANLPYQNVSETFL 356

QY 365 YPPAPILIHVSSEDTIEGFNIPRTIVILINGWGMQDPOLWMDATCFKPERF----- 419
DB 357 YPAAPLPLVRSPTEDIKVGIDVPRGTVMVWNAWAHTRDPLEWNEPEKPERFNGEGG 416

QY 420 VEGEE-KKLVAFGMRRACPGEPMAQSVSFTLGLLIQCFDMKRVSEEEKLDMTENNITPL 478
DB 417 GRGEDVHKLMPFGNGRRSCFGAGLGQKIVTLALGSLIQCFDMQKYNGEALDITETPGMAM 476

QY 479 SRLIPLEAMCKARPLATKI 497
DB 477 RKXIPLSALCQSRPIMSKL 495

```

RESULT 11

AAG50189
ID AAG50189 standard; Protein; 516 AA.

XX AC AAG50189;

XX DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 63574.

XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

XX Arabidopsis thaliana.
OS EP1033405-A2.
PN 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126284.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.


```
Db 372 YPAAPLLVPRSTEDIKVGIDVPRGTMVWNAWAHHRDPGLWNEPKFKPERFNGGEGG 431
Qy 420 VEGEE-KKLVAFGMGRACPGSPMAMQSVSFTLGLLIQCDFDKRVSEEEKLDMTENNWITL 478
Db 432 GRGEDVHKLMFGNGRRSCPGAGLGQKIYTLALGSLIQCDFDKQVNGEALDWTETPGWAM 491
Qy 479 SRLIPLEAMCKARPLATKI 497
Db 492 RKKIPLSALCQSRPIMSKL 510

RESULT 12
AAG44324
ID AAG44324 standard; Protein; 501 AA.
XX AC AAG44324;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 55508.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139432.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
```


PR	28-APR-1999;	99US-01314449.	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	99US-01320448.	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	99US-01320407.	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	99US-01322484.	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	99US-01322485.	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	99US-01322486.	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	99US-01322487.	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	99US-01322488.	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	99US-01322489.	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	99US-01322490.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-01322491.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-01322492.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-01322493.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-01322494.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-01322495.	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	99US-01322496.	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	99US-01322497.	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	99US-01322498.	PR	02-AUG-1999;	99US-0146389.
PR	24-MAY-1999;	99US-01322499.	PR	03-AUG-1999;	99US-0147038.
PR	25-MAY-1999;	99US-01322500.	PR	04-AUG-1999;	99US-0147204.
PR	27-MAY-1999;	99US-01322501.	PR	04-AUG-1999;	99US-0147302.
PR	28-MAY-1999;	99US-01322502.	PR	05-AUG-1999;	99US-0147192.
PR	01-JUN-1999;	99US-01322503.	PR	05-AUG-1999;	99US-0147260.
PR	03-JUN-1999;	99US-01322504.	PR	06-AUG-1999;	99US-0147303.
PR	04-JUN-1999;	99US-01322505.	PR	06-AUG-1999;	99US-0147416.
PR	07-JUN-1999;	99US-01322506.	PR	06-AUG-1999;	99US-0147493.
PR	08-JUN-1999;	99US-01322507.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-01322508.	PR	09-AUG-1999;	99US-0148171.
PR	10-JUN-1999;	99US-01322509.	PR	11-AUG-1999;	99US-0148319.
PR	14-JUN-1999;	99US-01322510.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-01322511.	PR	13-AUG-1999;	99US-0148565.
PR	16-JUN-1999;	99US-01322512.	PR	13-AUG-1999;	99US-0148684.
PR	17-JUN-1999;	99US-01322513.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-01322514.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-01322515.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-01322516.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-01322517.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-01322518.	PR	20-AUG-1999;	99US-0149829.
PR	18-JUN-1999;	99US-01322519.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-01322520.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-01322521.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-01322522.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-01322523.	PR	27-AUG-1999;	99US-0151065.
PR	21-JUN-1999;	99US-01322524.	PR	27-AUG-1999;	99US-0151066.
PR	22-JUN-1999;	99US-01322525.	PR	30-AUG-1999;	99US-0151080.
PR	23-JUN-1999;	99US-01322526.	PR	31-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-01322527.	PR	01-SEP-1999;	99US-0151438.
PR	24-JUN-1999;	99US-01322528.	PR	07-SEP-1999;	99US-0151930.
PR	28-JUN-1999;	99US-01322529.	PR	10-SEP-1999;	99US-0152363.
PR	29-JUN-1999;	99US-01322530.	PR	13-SEP-1999;	99US-0153070.
PR	30-JUN-1999;	99US-01322531.	PR	15-SEP-1999;	99US-0153758.
PR	01-JUL-1999;	99US-01322532.	PR	16-SEP-1999;	99US-0154018.
PR	02-JUL-1999;	99US-01322533.	PR	20-SEP-1999;	99US-0154039.
PR	08-JUL-1999;	99US-01322534.	PR	22-SEP-1999;	99US-0154779.
PR	08-JUL-1999;	99US-01322535.	PR	23-SEP-1999;	99US-0155139.
PR	09-JUL-1999;	99US-01322536.	PR	24-SEP-1999;	99US-0155486.
PR	12-JUL-1999;	99US-01322537.	PR	28-SEP-1999;	99US-0155659.
PR	13-JUL-1999;	99US-01322538.	PR	29-SEP-1999;	99US-0156458.
PR	14-JUL-1999;	99US-01322539.	PR	04-OCT-1999;	99US-0156596.
PR	15-JUL-1999;	99US-01322540.	PR	05-OCT-1999;	99US-0157117.
PR	16-JUL-1999;	99US-01322541.	PR	06-OCT-1999;	99US-0157753.
PR	16-JUL-1999;	99US-01322542.	PR	07-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-01322543.	PR	08-OCT-1999;	99US-0158029.
PR	19-JUL-1999;	99US-01322544.	PR	12-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-01322545.	PR	13-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-01322546.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-01322547.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-01322548.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-01322549.	PR	14-OCT-1999;	99US-0159329.
PR	20-JUL-1999;	99US-01322550.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-01322551.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-01322552.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-01322553.	PR	14-OCT-1999;	99US-0159638.

QY 357 IILETLRLYPAPILPHVSSDITIEGFNIPRDTIIVINGMQRDPOLMNDATCFKE 416
 Db 360 IVLETLRLHPATLLPHMASEDCKIGSYDMPRGTTLLVNAWAHEDPNTWDDPDSFKPE 419
 QY 417 RFVVEGEEKLVAFGMGRACPGEPNMAOSVFTLLGLIQCFWKVSEKLDMTENNWI 476
 Db 420 RFEKEBEAQKLLAFGLGRACPGSLAQRTVGLALGSLIQCFEWEVGNVEVDMKEGVGN 479
 QY 477 TLSRLIPLEAMCKARPLATKI 497
 Db 480 TVPKAIPLKAICKARPLHKI 500

RESULT 15
 AAG42047
 ID AAG42047 standard; Protein; 505 AA.
 XX AAG42047;
 AC XX
 DT 18-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 52390.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridization assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123348.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144633.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 20:34:14 ; Search time 69 Seconds

(without alignments)
1185.429 Million cell updates/sec

Title: US-09-868-546A-2

Perfect score: 2592

Sequence: 1 LVVVSYAVLFLVFLGVKF.....RLIPLEMCKARPLATKIGI 499

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 609560 seqs, 163917102 residues

Total number of hits satisfying chosen parameters: 609560

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	921.5	35.6	514	12	US-10-356-153-68
2	921.5	35.6	514	15	US-10-142-231-68
3	810.5	31.3	509	12	US-10-097-559-11
4	773	29.8	512	12	US-10-356-153-66
5	773	29.8	512	15	US-10-142-231-66
6	756	29.2	506	10	US-09-814-786-65
7	753	29.1	471	12	US-10-097-559-42
8	742	28.6	516	14	US-10-067-534-3
9	730	28.2	523	10	US-09-814-786-67
10	729	28.1	510	12	US-10-021-425-30
11	724.5	28.0	509	12	US-10-097-559-41
12	718	27.7	502	12	US-10-259-165-146
13	711.5	27.4	508	9	US-09-796-256A-2
14	707	27.3	520	14	US-10-067-534-4
15	704	27.2	508	9	US-09-931-267-4

16	702	27.1	508	9	US-09-931-267-5
17	695	26.8	325	12	US-10-259-165-102
18	695	26.8	325	12	US-10-259-165-434
19	678	26.2	542	12	US-10-289-757-107
20	673.5	26.0	510	10	US-09-814-786-66
21	671.5	25.9	477	9	US-09-931-267-6
22	669.5	25.8	477	9	US-09-931-267-7
23	666.5	25.7	543	12	US-10-289-757-183
24	664.5	25.6	532	12	US-10-289-757-106
25	657	25.3	526	12	US-10-361-460-73
26	657	25.3	529	16	US-10-174-693-405
27	652.5	25.2	535	14	US-10-067-534-2
28	651.5	25.1	514	10	US-09-947-027-4
29	651.5	25.1	514	14	US-10-091-009-4
30	645	24.9	519	12	US-10-097-559-9
31	638.5	24.6	531	11	US-09-899-642-2
32	631	24.3	495	12	US-10-021-425-29
33	618	23.8	511	9	US-09-796-256A-4
34	613	23.6	553	10	US-09-349-385-8
35	611	23.6	518	15	US-10-163-198-35
36	609.5	23.5	534	10	US-09-349-385-1
37	609.5	23.5	534	15	US-10-163-198-44
38	603.5	23.3	474	12	US-10-097-559-3
39	603	23.3	473	12	US-10-097-559-1
40	601	23.2	504	15	US-10-163-198-41
41	598	23.1	530	10	US-09-349-385-4
42	598	23.1	530	15	US-10-022-025A-3
43	598	23.1	530	15	US-10-163-198-46
44	596.5	23.0	516	15	US-10-163-198-23
45	580.5	22.4	514	15	US-10-163-198-97

ALIGNMENTS

RESULT 1

US-10-356-153-68
; Sequence 68, Application US/10356153
; Publication No. US20030166176A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: CYTOCHROME P450 OXYGENASES AND THEIR USES
; FILE REFERENCE: 62773
; CURRENT APPLICATION NUMBER: US/10/356,153
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: US/10/142,231
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 60/165,250
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 68
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Taxus cuspidata
US-10-356-153-68

Query Match 35.6%; Score 921.5; DB 12; Length 514;
Best Local Similarity 37.9%; Pred. No. 1.6e-71;
Matches 196; Conservative 103; Mismatches 175; Indels 43; Gaps 11;
QY 2 LVVVSYAVLFLVFLGVKFVFSRKLNN-----IPPGPPPLPIIGNLLEEQ--PIHR 52
DB 8 VILFTLALLLV-----VIORRRIRRHKKLVKQVAPQPSWFWLGNHLLLTQKVPIHR 60
QY 53 FFORMSKYQGNVSVLWFGSRSLAVWISSPTAYOECFTKHDVALANRLPSLS-GKVIYFNNT 111
DB 61 ILSSISESGPTMHLQGLRPAIVASSDLAKCEFTNDKAFASR-PRLSAGKHVGYDYK 119
QY 112 TVGSCSHGHWRNLRRTALDVLSTQRVHSFGIRSDETKRLMQRLVLAKNNEEFARV 171
DB 120 IFSMAFGYWNLRKMCCTIQLSATRIDSFRHVRVEVSALIRSLF---DSCOREDTPV 176


```
||||| : : : || : : : || : : :
Db 477 DLDMEIFGLSTPKPLATVIEPR-LSPKL 506

RESULT 4
US-10-356-153-66
; Sequence 66, Application US/10356153
; Publication No. US20030166176A1
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: CYTOCHROME P450 OXYGENASES AND THEIR USES
; FILE REFERENCE: 62773
; CURRENT APPLICATION NUMBER: US/10/356,153
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: US/10/142,231
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 60/165,250
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Taxus cuspidata
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(512)
; OTHER INFORMATION: "Xaa" equals any peptide
US-10-356-153-66.

Query Match 29.8%; Score 773; DB 12; Length 512;
Best Local Similarity 32.5%; Pred. No. 1.3e-58;
Matches 160; Conservative 125; Mismatches 178; Indels 30; Gaps 10;

QY 8 AVLFLVFLGVKVFQSRKLRN----IPGPPPLPIIGNLNLEQPIHRRFFORMSKOYGN 63
Db 22 AVLTIFLIGIFYL--RGLRNGRLPPGPPIWPVIGNLHQLGKLPNRLLELAKKHP 79

QY 64 VVSLWFGSLAVVISPTAYQECFTKHDVALANRLPSLSGKYIFYNNTTVGSCSHGEHWR 123
Db 80 IMLMGLGSPAVIVSSSAAKEVLKTHDLVFASRPESAAGKYIAYNKDIVFSPYGPYWR 139

QY 124 NLRRITADVLTSTQVHVSFGIRSDTKRLMORLVLAQNSNEEFARVEISSMFDLTYN 183
Db 140 QMKKICVVELLNARRIESLRVREEVSVII--RSVWEK--SKOGAVAVNLSKTLSSLTG 196

QY 184 NIMRMISGKRFYGESEMKNVBEAREFRTVTMELMGLANKGDHLPFLRWDFQNVK 243
Db 197 LMQIFSSNDGNS-----VTAIKEMSEVSETAGAFNIGDYFPWMDWMDLQGIOR 249

QY 244 RLKSIKRYDSTLNKILHEN--RASNDRO--NSMIDHLKLQETQPOYYTQIIGLAL 298
Db 250 RMTKAHDYFDQVITKIEHQHTRAMEDTQOPKDIIDALLQMENTDGVTTITMENIKAVVL 309

QY 299 AMLFGGTDSTGTLEWSLNLNHPVLKARDELDTQVGDRLNLES DLPKLPYLKII 358
Db 310 GIFLGAETSTTLEWAMSAMLENPEVAKVQEEIESVVGKRVKEMIWESMEYLQCVV 369

QY 359 LETLRLYPAPILIPHVSSEDITIEGFNIPRDTIVINGMQRDPQLMNDATCFKPERF 418
Db 370 KKTMRLYPAVPLIIPHESTQDCTVNGYFIPERTRILVNAWAIGKDPNVWDALAFKPRF 429

QY 419 ---DVEGEKK----LVAFGMRACGPEPMQSVFTLGLLIQCFDMKRVSEKLDMT 471
Db 430 LGXNVDLQKGEFFDMVFPFAGKGCPCGASMAVVTMEHALAQLMHCFOWR--IEGELDMS 487

QY 472 ENRWITLSRLIPL 484
Db 488 ERLAASVQKKVDL 500

RESULT 5
US-10-142-231-66
; Sequence 65, Application US/09814786
; Patent No. US20020100072A1
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yasuhiro
; KIKUCHI, Yasuhiro
; SHIMADA, Yukihisa
; OHBAYASHI, Masaya
; SHIMADA, Ritsuko
```



```

1 CHIKARA, Ritsuko
2 SHIMADA, Yaeushi
3 OKINAKA, Yaeushi
4 TITLE OF INVENTION: NOVEL PLANT GENES
5 NUMBER OF SEQUENCES: 67
6 CORRESPONDENCE ADDRESS:
7 ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
8 STREET: 30 Rockefeller Plaza
9 CITY: New York
10 STATE: New York
11 COUNTRY: U.S.A.
12 ZIP: 10112-3801
13 COMPUTER READABLE FORM:
14 MEDIA TYPE: Diskette - 3.50 inch, 720 Kb storage.
15 COMPUTER: IBM PS/V
16 OPERATING SYSTEM: MS-DOS Ver3.30
17 SOFTWARE: PATENT AID Ver1.0
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/09/814,786
20 FILING DATE: 23-Mar-2001
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: US/09/616,990
23 FILING DATE: 14-Jul-2000
24 APPLICATION NUMBER: JP44963/92
25 FILING DATE: 02-MAR-1992
26 ATTORNEY/AGENT INFORMATION:
27 NAME: Perry, Lawrence S.
28 REGISTRATION NUMBER: 31865
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: 212-218-2100
31 TELEFAX: 212-218-2200
32 INFORMATION FOR SEQ ID NO: 67:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 523 amino acids
35 TYPE: amino acid
36 STRANDEDNESS: single
37 TOPOLOGY: linear
38 MOLECULE TYPE: peptide
39 ORIGINAL SOURCE:
40 ORGANISM: Campanula medium
41 FEATURE:
42 NAME/KEY: CDS
43 LOCATION: 180 to 1748
44 IDENTIFICATION METHOD: by experiment
45 SEQUENCE DESCRIPTION: SEQ ID NO: 67
46
47 US-09-814-786-67
48
49 Query Match 28.2%; Score 730; DB 10; Length 523;
50 Best Local Similarity 31.3%; Pred. No. 7.5e-55;
51 Matches 159; Conservative 117; Mismatches 208; Indels 24; Gaps 7;
52
53 QY 4 VVSYAVLFLVFLGVGVFQSKRLRNIPGPPPLPIIGNLNLEQPIHRRFFORMSKOVGN 63
54 12 LVAAISLYLATYSFIRLFPKPSHHHLPPTGCTGWPILGALPLIGTDPHVSILADWAVKGP 71
55
56 QY 64 VVSLWFGSRLVAVISSPTAYQECFTKHDVALANRLPSLSKGYFYNNNTTVGSCSHGHR 123
57 12 IMYKLGSKTGVASNPKAARAFKTHDANFNRPIDGGPTLYLAYNAQDMVFAEYGPWK 131
58
59 QY 124 NLRITALDVLSTQRVHSPSGIRSDETKRLMQRLVLAKNSEBEFARVEISSMFDLTYN 183
60 132 LLKCLGSLHMLGPKALEDAHVAVKSEVGHMLKEMY--EQSSKSVFVVPVPEMLTYAMAN 189
61
62 QY 184 NINRMISGKRFYGESEMKVBEA---REFRETVTETMLELMLGLANKGDHLPLFLRWFDPQ 239
63 190 MIGRIILSRPFFVITSLKDSASASASVSEFYVMVLMRMAGFLIGDIFPIYANMDLQ 249
64
65 QY 240 NVEKRLKISKRYDSINILKHEN--RASNDRONSMIDHLKLOETQPYQYTDQI----- 292
66 250 GIQRDMKVIQKFDVLLNKKMIKEHTESAHRKDN--PDFDLILMAAQTENGTGQIQLNVN 307
67
68 QY 293 IKGIALAMLPGGTDSSTGTLWLSLNLNHPVLKARDELDTQVGDRLNNSDLPLP 352
69 352
70

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Db	190	MICRILSRPPVITSEKLSDSASASASVSFEQYVMWELMERMAGLFNIGCDRPIYIAWDLQ	249
Qy	240	NVEKRLKISKRYSDSLNLKILHEN-RASNDQRNSMIDHLKLOETQOYYTQDI-----	292
Db	250	GIQDMKVIQKFDVLLNKMIKEHTESAHRKDN--PDFDLILMAAQTENTEGIQLLNVN	307
Qy	293	IKGLALAMLFPGGTDSSTGTLWELSNLNLNHPVILKARDELDTVQGDRLINESDLPKLP	352

Db 308 VKALLDLFTAGTTSSTVIEWALAEMLNHRQILNRAHEMDQVIGNRRLEQSDIPNLP 367
Qy 353 YLRKIILETLRLYPAPILPHVSSSEDTITBGFNIPRDTIVINGMGMQDQPOLWMDATC 412
Db 368 YFOAICKETFKHSTPLNLPRISTEACEVDGPHIPKNTLIVNIAIGDRPKWENPLD 427
Qy 413 FKPERF-----DVEGEKKLVAFGGRACGPEPMQSVSFTLGLLIQCFDKRV - 463
Db 428 FTERFLSEKHAIDPRGNHPELIPFGAGRRICAGARMGAASVEYILGTLVHFSDFWKLPD 487
Qy 464 SEEKLDMTENNWTLSRLIPLEAMCKAR 491
Db 488 GWVEVNMEEFGIALQKKVPLSAIVTPR 515
RESULT 10
US-10-021-425-30
; Sequence 30, Application US/10021425
; Publication No. US20030148420A1
; GENERAL INFORMATION:
; APPLICANT: Suzanne L. Bolten
; APPLICANT: Alan M. Easton
; APPLICANT: Leslie C. Engel
; APPLICANT: Dean M. Messing
; APPLICANT: John S. Ng
; APPLICANT: Beverly A. Reitz
; APPLICANT: Scott A. Vaccaro
; APPLICANT: Mark C. Walker
; APPLICANT: Ping T. Wang
; APPLICANT: Robin A. Weinberg..
; TITLE OF INVENTION: Aspergillus ochraceus 11 alpha
; FILE REFERENCE: S03196-00-US
; CURRENT APPLICATION NUMBER: US/10/021,425
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: USSN 60/244,300
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Glycine max AAB94588
US-10-021-425-30
Query Match 28.1%; Score 729; DB 12; Length 510;
Best Local Similarity 31.1%; Pred. No. 8,8e-55;
Matches 158; Conservative 116; Mismatches 192; Indels 42; Gaps 10;
Qy 7 YAVLFLVLFLGVKVFV-----QSRKLRNIPPPPPPLPIIGNLNL--EQPIHFFQ 55
Db 11 PSIVFITSILFFVFFKLVQSRDSKTSSTCKLPPGPRTLPLIGNIHQIVGSLPVHYLK 70
Qy 56 RMSQYGNVSLWFGSLAVVISPTAYQECFTKHDVALANRLPSLSGKIYFYNNTVGS 115
Db 71 NLADYQFIMHLKLGVEVNIITVPEMAQEIIMTKHDNLNFSRDPVLSRVYNGSGIVF 130
Qy 116 CSHEGHRNLRITDVLSTQVHSGFSGISDETKLMORLVLAQNSNEEFARVEISS 175
Db 131 SOHGDYWLQKICTVELLTAKRVQSFRIEEVAELVKK--TAATASEGGSFNLTQ 188
Qy 176 MFNDLYNNIMRISGKRYGEESEMKVVEARFRETVTBMLGLANKGDHLPFLRW 235
Db 189 SIYSMTFG-----IAARAFAFKKRYQQV-----FISNMHQLMLLGGFSVADLYPSRV 238
Qy 236 FDFQNVKRLKSIKRYDSILNKLHE-----NRASNDQ--NSMIDHLKLOETQPYT 289
Db 239 FOMGATGKLEKVRVTORVLODIIDEHNRNRSSEAREVEDLVDLKFKQKSEFRLT 298
Qy 230 DQI1KGLALAMFGGTDSTGTLEWSLNLNHPVLKARDELDTQVQDRLNESDLP 349
Db 299 DDNIKAVIQDIFIGGETSSSVVWGMSELINPRVMEAAQVRRVYDSKGYVDETELH 358

Qy 350 KLPYLRKIILETLRLYPAPILPHVSSSEDTITBGFNIPRDTIVINGMGMQDQPOLWMD 409
Db 359 QLIYLSKIIKETMRLHPPVPLVPRVSRERCOINGVEIPSKTRIIINAWAIGRNPKYGCE 418
Qy 410 ATCFKPERF-----DVEGEKKLVAFGGRACGPEPMQSVSFTLGLLIQCFDWK--- 461
Db 419 TSEFKPERFLNSSIDFRGTDPEFIPFGAGRRICPGITFAIPNIELPLAQLLYHFDWKLPN 478
Qy 462 RVSEKKLDMTENNWTLSR-----LIPL 484
Db 479 KMKNEHLDMTESNGITLRRQNDCLIP 506
RESULT 11
US-10-097-559-41
; Sequence 41, Application US/10097559
; Publication No. US20030166255A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: Ralston, Lyle P.
; TITLE OF INVENTION: Cytochrome P450s and Uses Thereof
; FILE REFERENCE: 07678/100003
; CURRENT APPLICATION NUMBER: US/10/097,559
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: US 60/274,241
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 60/275,597
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Nepeta racemosa
US-10-097-559-41
Query Match 28.0%; Score 724.5; DB 12; Length 509;
Best Local Similarity 32.4%; Pred. No. 2.2e-54;
Matches 164; Conservative 101; Mismatches 208; Indels 33; Gaps 10;
Qy 4 VVSYAVLFLVLFLGVKVFVQSRKLRNI-----PCGPPPLPIIGNLNLLEQPIHFFQ 57
Db 1 MVSLSYFLIALCTLPFLFLNKRWSYSGKTPPPPKLPVIGNUHQGLYPHRYLQSL 60
Qy 58 SKQYGNVSLWFGSLAVVISPTAYQECFTKHDVALANRLPSLS-GKIYFYNNTVGSC 116
Db 61 SRRYGLMQLHGFSGVPLVASSPEAREIMKQDIVFSNR-PKMSIANRLFENNRDVAFT 119
Qy 117 SHGHEWRNLRITDVLSTQVHSGFSGISDETKLMORLVLAQNSNEEFARVEISS 176
Db 120 QYGEYWRQIRSCVQLQLSNKRVSPPRVREETSIMVEKIMQLGSSSS---TPVNLSEL 176
Qy 177 FNDLTNNIMRISGKRYGEESEMKVVEARFRETVTBMLGLANKGDHLPFLRW 236
Db 177 LLSLTNDVVCVRVLGKKGKGGNGS-----BEVDKLEMLTEIQNLMGISPWEIFPWNWT 232
Qy 237 -DFQNVKRLKSIKRYDSILNKLHENR-----ASNDQRNSMIDHLKLOETQPY - 287
Db 233 RFPDGVQVDRIVKAFDGFLESVIEKHERDQKDGDDGALDFVDILLQFORENQRS 292
Qy 288 -YTDQI1KGLALAMFGGTDSTGTLEWSLNLNHPVLKARDELDTQVQDRLNES 346
Db 293 PVEDDTVKALILDPMFVAGTDTTATALEWAEVLKPNRKLQNEVREVAGSKAETEE 352
Qy 347 DLPLKPLVLRKIILETLRLYPAPILPHVSSSEDTITBGFNIPRDTIVINGMGMQDQPOL 406
Db 353 DLEKMYLKASIKESURLHVPVLLVPRESTRTNVLGYDIASGTRVLINAWAIARDPSV 412
Qy 407 WNDATCFKPERF-----DVEGEKKLVAFGGRACGPEPMQSVSFTLGLLIQCFDW - 460
Db 413 WENPEFLEPERFLDSSIDYKGLHPELLPFGAGRGCGPGATFAVAIDELALAKLVHFDG 472
Qy 461 --KRVSEKKLDMTENNWTLSRLIPL 484

US-10-067-534-4
; Sequence 4, Application US/10067534
; Publication No. US20020187538A1
; GENERAL INFORMATION:
; APPLICANT: Essenberg, Margaret K.
; APPLICANT: Chen, Xiao-Ya
; APPLICANT: Luo, Ping
; APPLICANT: Wang, Yan-Hong
; TITLE OF INVENTION: cDNA Clone of (+)-Delta-Cadinene-8-Hydroxylase Gene from Cotton
; FILE REFERENCE: 006602-113
; CURRENT APPLICATION NUMBER: US/10/067,534
; CURRENT FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: US 60/267,160
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-067-534-4

Query Match 27.3%; Score 707; DB 14; Length 520;
Best Local Similarity 31.2%; Pred. No. 7.4e-53;
Matches 159; Conservative 108; Mismatches 204; Indels 38; Gaps 10;
QY 2 LVVVSAYAVLFLVFLGVK-FVFSQKRLNIPPGPPPLPIIGNLNLLEQPIHRFFORMSKQ 60
DB 15 LTPYAVAVLITATFSILWYIFKSPQPLPPGRGLPIVGNLPLDPDLHTYETKLAQS 74
QY 61 YGNVVSWMFSGRLAVVSSPTAYQECFTKHDVALANRLPSLSGKYIFYNNTTVGSCSHGE 120
DB 75 HGPIFKINGSLKLVVWNSPGLSEILKQDINFSDVPLTARAVTYGGLDLVWLPYGA 134
QY 121 HWRNLRRITADVLTSTQVHSFGIRSDETK--RLMORLVAK--NSNEEEFARVEIS 174
DB 135 EWRMLRVCAAKLFSRKTLDSTFYELRRKEIRERTRCLYOKGLEKSPVNVGSQLF----- 188
QY 175 SMFNDLTNNI--MRMISGRFYGESEEMKNVEEAREFRETVTMEMLMGLANKGDHLPFL 233
DB 189 -----LTMNLMNMLWGSVKAEDEME---SVGTPEKGVISEITRLLGVENVSDFFPML 239
QY 234 RWFDFQNVKRLKLSIKRYDSILNKILHE-----NRASNDQ--NSMDHLLKLOETPOQY 287
DB 240 ARFDLQGLVKKVHLYARDLDAILDRAIEQMORLSRSDGDDGECKDFLQHLMKLRDQEADS 299
QY 288 --YTDQIIKGLALAMLFGGTDSSTGTLEWSLSNLLNHPVLKKADELDTQVGDRLN 344
DB 300 DVPITMNVKAVLMDVMVGGTSESTNTIEFVMAELISNPELMRAAQDELDEVGKDIVE 359
QY 345 ESDLPKLPYLRIKILETLRLYPAPILIPHVSSDITIEGFIPIHVSSEDIIEGFIPIHVSSEDI 404
DB 360 ESHITSLPVLAVLKSTLRLYPTIPLVPHRSETALVGGVTIPKNTKIFINVNSIQRP 419
QY 405 QLWNDATCFKPERF-----DVEGEKKLVAFGMRRACPGEPMAQSVSTLGLLQCF 458
DB 420 NVWEYPTFERPERFLDKSCDFTGTDYSYLPFGSGRRICAGTALAERMILYTLTLHSF 479
QY 459 DWKRVSEKLDMTENNNTLSRLIPEAM 487
DB 480 DWTIPDGHVLDLEKFGVILKTKPLVAL 508

RESULT 15
US-09-931-267-4
; Sequence 4, Application US/09931267
; Patent No. US20020062496A1
; GENERAL INFORMATION:
; APPLICANT: Chapple, Clinton
; APPLICANT: Franke, Rochus
; APPLICANT: Ruegger, Maxwell
; TITLE OF INVENTION: Genes Encoding P-Coumarate 3-Hydroxylase (C3H) and Methods of Use
; FILE REFERENCE: N1422-005

; CURRENT APPLICATION NUMBER: US/09/931,267
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 60/225554
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-931-267-4
Query Match 27.2%; Score 704; DB 9; Length 508;
Best Local Similarity 30.9%; Pred. No. 1.3e-52;
Matches 156; Conservative 113; Mismatches 202; Indels 34; Gaps 9;
QY 1 LLVVVSAYAVLFLVFLGVKVFQSRKLR-NIPPGPPPLPIIGNLNLLEQPIHRFFORMSK 59
DB 11 IAAVWSYKLI-----QRLRYKFFPGSPKPIVGNLYDIKPVRFRCYIYEAQ 56
QY 60 QYGNVVSWMFSGRLAVVSSPTAYQECFTKHDVALANRLPSLSGKYIFYNNTTVGSCSHG 119
DB 57 SYGPIISVMIGSLNVVSSAELAKEVLEKHDOKLADRHRNRSTFASRNGQDLIWDYD 116
QY 120 EWRNLRRITADVLTSTQVHSFGIRSDETKRLMORLVAKNSNEEEFARVEISSMFD 179
DB 117 PHYVKRVKVTLELFTPKLESIRPIREDEVAMVESVFRDCNLNPNRAKGLQRLKYLGA 176
QY 180 LTVNNIMRMISGRFYGESEEMKNVEEAREFRETVTMEMLMGLANKGDHLPFLRWFDFQ 239
DB 177 VAFNNITRIAFGRFNNAGVVD--EGLFKAIVSNGKLKGLASLSIAEHIPWRWM-FP 233
QY 240 NVEKRLKSISKRYDSILNKILHEN-----RASNDQNSMIDHLLKLOETPOQY-YTDQIIK 294
DB 234 ADEKAFAEHAGARRDLITRAIMEEHTLAROKSSGAKQHFVDALITLKD---QYDLS EDTII 290
QY 295 GLALAMLFGGTDSSTGTLEWSLSNLLNHPVLKKADELDTQVGDRLNLESDLPKLPYL 354
DB 291 GLLWDMITAGMDTTAITAEWAMAEIMKNPRVQKQVEEFDREVGVGLDRILTEADFSRLPYL 350
QY 355 RKIILETLRLYPAPILIPHVSSDITIEGFIPIHVSSEDIIEGFIPIHVSSEDIIEGFIPIHVSSEDI 414
DB 351 QCVVKSFRLLHPTPLMLPHRSNADVKIGYDIPKGSNVHVNVAWVARDPAVWKNPEFR 410
QY 415 PERF-----DVEGEKKLVAFGMRRACPGEPMAQSVSTLGLLQCFDW---KRVSEE 466
DB 411 PERFLEEDVDMKGDHDFLLPFGAGRRVCPGAQLGNLVTSMMSHLLHFWVTPPQGTKE 470
QY 467 KLDMTENNNTLSRLIPEAMCKAR 491
DB 471 EIDMSENPGLVVTYMRTPVOAVTPR 495

Search completed: October 20, 2003, 20:44:04
Job time : 74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 20, 2003, 20:30:09 ; Search time 30 Seconds
(without alignments)
703.771 Million cell updates/sec

Title: US-09-868-546A-2
Perfect score: 2592
Sequence: 1 LLVVSVAVLFLVFLGVKF.....RLIPLEAMCKARPLATKIGI 499

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778.5	30.0	513	4	US-09-351-229-4
2	764.5	29.5	532	3	US-08-948-564-10
3	756	29.2	506	1	US-08-313-075A-38
4	756	29.2	506	3	US-08-606-505B-65
5	756	29.2	506	3	US-09-616-990-65
6	747	28.8	471	4	US-09-126-420A-20
7	731.5	28.2	508	4	US-09-126-420A-19
8	730	28.2	523	3	US-08-606-505B-67
9	730	28.2	523	3	US-09-616-990-67
10	729	28.1	510	3	US-08-948-564-4
11	725.5	28.0	495	1	US-08-532-065B-2
12	713	27.5	490	4	US-09-126-420A-16
13	711.5	27.4	508	3	US-08-991-677-2
14	684.5	26.4	509	3	US-08-948-564-18
15	681.5	26.3	496	1	US-08-313-075A-50
16	673.5	26.0	510	3	US-08-606-505B-66
17	673.5	26.0	510	3	US-09-616-990-66
18	665.5	25.7	513	3	US-08-948-564-2
19	664	25.6	504	4	US-09-126-420A-18
20	662.5	25.6	513	4	US-09-564-808-6
21	660	25.5	520	2	US-09-091-432-2
22	660	25.5	520	2	US-09-387-663-2
23	657	25.3	529	4	US-09-615-192A-405
24	648	25.0	520	4	US-09-564-808-2
25	648	25.0	520	4	US-09-564-808-4
26	638.5	24.6	531	4	US-09-380-420C-2
27	633	24.4	516	3	US-08-948-564-12

28	621	24.0	504	4	US-09-499-302A-9	Sequence 9, Appli
29	618	23.8	511	3	US-08-991-677-4	Sequence 4, Appli
30	607	23.4	521	3	US-08-948-564-14	Sequence 14, Appli
31	598.5	23.1	509	4	US-09-499-302A-10	Sequence 10, Appli
32	597.5	23.1	502	4	US-09-499-302A-7	Sequence 7, Appli
33	591.5	22.8	476	1	US-08-313-075A-30	Sequence 30, Appli
34	584	22.5	502	4	US-09-499-302A-2	Sequence 2, Appli
35	583	22.5	504	4	US-09-499-302A-6	Sequence 6, Appli
36	578.5	22.3	496	3	US-09-292-768-64	Sequence 64, Appli
37	577.5	22.3	496	3	US-09-292-768-2	Sequence 2, Appli
38	577.5	22.3	496	3	US-09-292-768-66	Sequence 66, Appli
39	577.5	22.3	500	3	US-09-292-768-68	Sequence 68, Appli
40	576.5	22.2	500	3	US-09-292-768-4	Sequence 4, Appli
41	575.5	22.2	500	3	US-09-292-768-70	Sequence 70, Appli
42	574.5	22.2	509	4	US-09-499-302A-8	Sequence 8, Appli
43	570.5	22.0	498	3	US-09-292-768-0	Sequence 6, Appli
44	567.5	21.9	496	3	US-08-881-784-1	Sequence 1, Appli
45	564	21.8	523	3	US-08-948-564-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-351-229-4
; Sequence 4, Application US/09351229
; Patent No. 6380465
; GENERAL INFORMATION:
; APPLICANT: Barrett, Michael
; TITLE OF INVENTION: Cytochrome P450 Enzymes and Related Compounds and
; FILE REFERENCE: P-1049
; CURRENT APPLICATION NUMBER: US/09/351,229
; CURRENT FILING DATE: 1999-07-11
; EARLIER APPLICATION NUMBER: 60/092,596
; EARLIER FILING DATE: 1998-07-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Zea mays
US-09-351-229-4

Query Match	30.0%;	Score	778.5;	DB	4;	Length	513;
Best Local Similarity	35.3%;	Pred. No.	9.3e-64;				
Matches	182;	Conservative	99;	Mismatches	181;	Indels	53;
						Gaps	15;
Qy	15	FLGVK-----PVFQSRK-LRNIPGPPPLPIIGNLNLLLEQPIHRRFFORMSKOYGNVVS	66				
Db	5	FLGIALCAAAALFLRGRPPVYVPPGPKPWPPIIGNLNLMGELPHRSMNELSKRYGPDQM	64				
Qy	67	LWFGSLAVVISPTAYQECFTK-HDVALANRLPFLSGKIFYFYNTTGVSCSHGEHW---	122				
Db	65	LWFGS-LPVVVGASAEAKLFLATNDAAFSRFAVGKYTAYD-----CS-GLLWAPF	116				
Qy	123	-----RMLRRITADVLSSTORVHSFGIRSDETKRLMORVLAKNSNEEERARVEISSMFN	178				
Db	117	EPYLRQARRICATELFSATRLSEFEHIDEVRMLRQLRQAAGRT-----VLRLDYQL	170				
Qy	179	DLTYNIMRMISKRFYGESEMKVNEEA-----REFRETVMLEMLGLANKGDHLPFL	233				
Db	171	MLALGVISRVLQKKYVMEEAADGEGDAPAITPAEFREVMDFEALHGFAGNIDYIPWL	230				
Qy	234	RWDFQNVKRLKLSISKRYDSILNKIL--HENRASNDRON-----SMIDHLLKLOE---TQ	284				
Db	231	DWLDLQGYVARMKMKARFCRFLERVLDVNERRLREGGNFVAKMDLVLQLADDTSL	290				
Qy	285	PQYTDQIKGLALAMFGGTDSTGTLEWSLNLNHPDEVLKARDELDTQGVQDRLN	344				
Db	291	VQLSRDN-VKAITQDILLIAGTDSNANTLEWAVSELLKNPKILAKAMELNHVRPDLVT	349				
Qy	345	ESDLPKPLRKLILETLRLYPAPILIPHVSSDITIEGFNIPRODTIVINGMGWRDP	404				

```

Db 350 ESDLPRLPYIEAVLKETMRVHPAAPMLAPHVAREDTSDVGYDVLAGTVLPFINVWAIGRDP 409
Qy 405 QLWNDATCFKPERF-----DVEGEKKIVAFGMGRRACPCPEWAMQSVSFTLGLLIQCFD 459
Db 410 GLWDAPEERPRFVESKIDVRGHDFQLPFGSGRRMCPGINLAKVMAUSLANLUHGFE 469
Qy 460 WK---RVSEKLDMTENNWITLSRLIPLEAMCKAR 491
Db 470 WRLPDGVTAELSMDEAFKLAVPRKFPLMVVAEPR 504

RESULT 2
US-08-948-564-10
; Sequence 10, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
; APPLICANT: Siminszky, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Corbin, Frederick T.
; TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
; TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia C. Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 6121512th Carolina
; COUNTRY: USA
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400
; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-948-564-10

```

```

237 DFNQVEKRLKSIKRYDYSIINKILHE-----NRASNDRONSMIDHLK-LQETQ-POY 289
253 DINGEYKAMRTASELDPLVEGLEEHRKRAFNDAKEEQDNFMDVNLVNLKDAEISGY 312
288 YTDQIIKGLALAMLFGGTSDSGTGLEWSLNLNHPEVLKKARFDELDQVGQDRLLNESD 347
313 DSDTIKATCNLIILASDITMTISLTWVLSLLLNHOMELKKVQDELDITYIGKDRKVESD 372
348 LPLKPYLRKIILETRLRYPAPILPHVSSDITIE-GFNIIPRDTIVIIINGMGORDPOL 406
373 ITKLVLQAIWKTMRLYPPSPPLITLRAAMEDCTFSGGYHIPAGTRLMVNAWKIHRDGRV 432
407 WNDATCKPERF-----DVEGEKKLVAFGMGRACPGPEPMAMQSVSFTLGLLIQCFD 459
433 WSDPHDPKGRFLTSKDKVDVKGNQYELVPPGSGRRACPGASLALRVVHLTWRLHSFN 492
460 WKRVSEKLDWNTENNWTLSRLPLPLEAMCKARPLATKI 497
493 VAPSNQVDMVTESIGLTNLKATPLEILLTFR-LDTKL 529

RESULT 3
US-08-313-075A-38
; Sequence 38, Application US/08313075A
; Patent No. 5639870
; GENERAL INFORMATION:
; APPLICANT: Holton, Timothy A.
; APPLICANT: Cornish, Edwina C.
; APPLICANT: Tanaka, Yoshikazu
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID
; TITLE OF INVENTION: PATHWAY ENZYMES AND USES THEREFOR
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/313,075A
; APPLICATION NUMBER: AU PL 1538/92
; FILING DATE: 27-MAR-1992
; PRIORITY INFORMATION:
; APPLICATION NUMBER: AU PL 6698/93
; FILING DATE: 07-JAN-1993
; PRIORITY INFORMATION:
; APPLICATION NUMBER: AU PCT/AU93/00127
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9433
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-313-075A-38

```

Query Match	29.2%;	Score 756;	DB 1;	Length 506;
Best Local Similarity	32.1%;	Pred. No. 1.1e-61;		
Matches 159;	Conservative 114;	Mismatches 200;	Indels 22;	Gaps 5
QY	10	LFLVFLGVKFVFSQRKLNIPTGPPPLPIIIGNLILLEQIHRFFQFMSKQYGVVSLMF	69	
DB	13	IFLIAHIISLTIISKTGGRHUPPGRWPFVIGALPLIGAMPVSLAKMAKKYGAIMYLKV	72	
QY	70	GSRLAVIISSTPAYQECPTKHDVALANRLPSLSGKIIFYNNNTTVGSCSHGEHNRNLRIT	129	
DB	73	GTCGNVASTPDAKAFKLTLIDINFSRPPNAGATHLAYNAQDMVFAHYGPRKKLRKLS	132	
QY	130	ALDVLSUTQRVHSFGIRSDYKRLMQRLVLAKNSNEEFARVETISSMFDLTVNNIMRM	189	
DB	133	NLHMGGKALENWANVRANELGHMLKSM---SDMSREGQRVVVAEMLTAFAMAMTIGQM	243	
QY	190	SKCRFYGESEMKQVVEAREFRETVTMEMLMGLANKGDHLPFLRWFDQFQVEKRLKLSIS	249	
DB	189	LSKRVFVDKG-----VENEFPKMVVELMTIAGYFNIGDFIPCLAMWMDQIGIEKKRKLIH	243	
QY	250	KRYDSILNKILHENRASDRQNSMIDHLLKQE---TPQYVTDQIIKGLALAMLFGGT	305	
DB	244	KKFDALLTKMDEHKATTYERKGPDPFLDVVMGNDSEGERLSTTNIKALLNLFTAGT	303	
QY	306	DSSTGTLBWSJNLNLNHPVLKKADELDTQVGDRLNLSDDPKLPYLKRIILETLRIY	365	
DB	304	DTSSSAIEWALAAEMKNPAILKKAQAEQDVIGNRNRLLESIDIPNLPYLRAICKETFRKH	363	
QY	366	PAPLILPHVASEDITIEGFNIPRDTIVINGWQMRDQPLWNDATCFKPERF-----	418	
DB	364	PSTPLNLRISNEPCIVDGYIYPKNRLSVNINWAIGRDPQVWENPUEFNFPERFLSGRNSK	423	
QY	419	-DVEGEKKLVAFGMGRACFCEPQWAMQSVFTLGLLIQCDFDKWRKVSSE--KLDMTENNMI	476	
DB	424	IDPRGNDPELFPFGNGRIICAGTNGIYVWVEIYLGTLVHSFDDKVLPSVEVLNWEAFGL	483	
QY	477	TLISRLIIPLEAMCKAR	491	
DB	484	ALOKAVPLEAMVTPR	498	

RESULT 4

US-08-606-505B-65

; Sequence 65, Application US/08606505B

; Patent No. 6114601

; GENERAL INFORMATION:

; APPLICANT: KIKUCHI, Yasuhiro

; APPLICANT: KIYOKAWA, Shigeto

; APPLICANT: SHIMADA, Yukihiisa

; APPLICANT: OHBAYASHI, Masaya

; APPLICANT: SHIMADA, Ritsuko

; APPLICANT: OKINAKA, Yasushi

; TITLE OF INVENTION: NOVEL PLANT GENES

; NUMBER OF SEQUENCES: 67

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10112-3801

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.

; COMPUTER: IBM PS/V

; OPERATING SYSTEM: MS-DOS Ver3.30

; SOFTWARE: PATENT AID Ver1.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/606,505B

; FILING DATE: 23-MAR-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP44963/92

; FILING DATE: 02-MAR-1992

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Perry, Lawrence S.
; REGISTRATION NUMBER: 31865
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-218-2100
; TELEFAX: 212-218-2200
; INFORMATION FOR SEQ ID NO: 65 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 506 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Petunia hybrida
; STRAIN: Falcon Blue
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 116 to 1633
; IDENTIFICATION METHOD: by experiment
; US -08-606-505B-65

Query Match          29.2%; Score 756; DB 3; Length 506;
Best Local Similarity 32.1%; Pred. No. 1.le-61;
Matches 159; Conservative 114; Mismatches 200; Indels 22; Gaps 5

Qy 10 LFLVLFGLGVKVFQSRKLRIIPGPPPLPIIGNLNLEQPIHRRPFQMSKOYGNVSLWFP 69
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 13 IFLIAHHIITLISKTTGRHLPGPRGPVIGALPLLGAHPHVSLAKNAKKYCAIMTLKV 72
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 70 GSRLAVVISPTAYQCFTKHVALANRSLSGKYIFYNTTVGSCSHGEHWENLRARIT 129
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 73 GTCGMAVASTPDAAKAFLLKTLDINFNRPNAGATHLAYNAQDMVFPAHYGPRWKLLRKLS 132
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 130 ALDVLSTQRVHSFGSIRDETETKLMOEVLVLAKNSEEFARVEISSMFNDLTNNINRMFI 189
Db   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 133 NLHWLGKALLENWANVPANELGHMLKSM----SDMSEGOQVVVAEMLTTFAMANNIGQVM 188
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 190 SGKFYGEESMKNVBSAREPRTVTTEMLBLMLGLANKGDHLPLFRWFDFQNVKRLKSIS 249
Db   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 189 LSKRVFVDKG-----VEVNEFKDMVVELMTIAGYFNIGDTIPCILAWMDLOQIEKRMRKLH 243
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 250 KRYDSIIINKILHNRAANDRONSKI DILLKLOE-----TOPOYYTDQI I KGLALAMLFGGT 305
Db   ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 244 KKFDALLTKNFDEHKATTYERKGPDPDLVVMENGDNSEGERLSTTNIKALLLNLTAGT 303
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 306 DSSGTGLEWSLNLNHPEVLKKARDDELDTQVGODRLINESDLPKLPVKRIILLETIRLY 365
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 304 DTSSAIAEWALAEMKKNPAILKKAQAEMDQVIGENRRLESDDIENLPYLRAICKETPKKH 363
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 366 PPAPILPHVSSEDITIEGFNI PRDTIVINGMGORDPOLWN DATCFKPERF----- 418
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 364 PSTPLNLPRI SNPCIVDGYI PKNTRLSVNIWAIGRDQPWVENPLEFNFPERFILSGRSNK 423
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 419 -DVEGEKKLVACGMGRACPGEPMAQOSVSFTGLLI QCFDWKRVSEE-KLDMTENNWII 476
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 424 IDPRGNDFFELIPGAGRRI CAGTRMGIVMWVEYITGLTVHGSFWDKWL PSEVIELNMEEAFGL 483
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 477 TLSLELIPLEAMCKRAR 491
Db   | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 484 ALOKAVPLEAMVTPR 498
Db   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
US-09-616-990-65
; Sequence 65, Application US/09616990
; Patent No. 6232109
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yasuhiro
; ; KIKOKAWA, Shigeto
; ; SHIWADA, Yukihisa
; ; OHBAYASHI, Masaya
; ; SHIWADA, Ritsuko
```

[illegible]

Db 412 EEFLPERVNNVDFKQDFQLIPFGAGRCPCGIAFGISSVEISLANILLYFNWE 467

RESULT 7
US-09-126-420A-19
; Sequence 19, Application US/09126420A
; Patent No. 6176753
; GENERAL INFORMATION:
; APPLICANT: BATAARD, YANNICK
; APPLICANT: ROBINEAU, TIBURCE
; APPLICANT: DURST, FRANCIS
; APPLICANT: WERCK-REICHAUT, DANIELE
; APPLICANT: DIDIERJEAN, LUC
; TITLE OF INVENTION: PURIFIED CYTOCHROME P450 CYP76B1 FROM HELIANTHUS
; TITLE OF INVENTION: TUBEROSUS AND ITS APPLICATIONS AS BIOCATALYST IN
; TITLE OF INVENTION: PARTICULAR FOR THE DEGRADATION OF ENVIRONMENTAL
; TITLE OF INVENTION: POLLUTANTS AND FOR ALTERING THE RESISTANCE OF PLANTS S
; FILE REFERENCE: 03715.0032
; CURRENT APPLICATION NUMBER: US/09/126.420A
; CURRENT FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/054,351
; PRIOR FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 19
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Petunia x hybrida
US-09-126-420A-19

Query Match 28.2%; Score 731.5; DB 4; Length 508;
Best Local Similarity 31.4%; Pred. No. 2.1e-59;
Matches 159; Conservative 112; Mismatches 206; Indels 29; Gaps 6;

Qy 9 VFLVFLGVKVFQSRKLRNIPGPPPLPIIGNLNLEQIHRFFORMSKQYGVNLSW 68
Db 12 LIFLTHIFITLSITNGRLPGPGWFGVIGALPILGAMPHVSLAKMAKYGAINYLK 71

Qy 69 FGSRLAVISSPTAYQBCFTKHDVALANRLPSLSKGKIFYNNTTVGSCSGHEHWRNLRRI 128
Db 72 VGTGCMVAVSTPAADAKAFKTLDFNFSNRPNAGATHLAYGAQDMVFARYGPRWKLRLK 131

Qy 129 TALDVLSTQVHVSFGIRSDTKLMQRLVLAKNSNEEFARVEISSMFNDLTNNMRM 188
Db 132 SNLHMGKALENANVRANLGHMLKSMFDMSEGE----RVVVAEMLTFAMANTIGQV 187

Qy 189 ISGRKRYFGESEEMKNVBEAREFRETVTETMLMGLMGLANKGDHLPFLRWFDONVEKRLSI 248
Db 188 ILSKRVF-----VKNKGVNEFKDMVVELMTTAGYFNIGDFIPCLAMNDLOGIEKMKRL 242

Qy 249 SKRYDSIILNKILHNENRASNDRQNSMIDHLLKLQETQPOYYTDQI-----IKGLALAMLFGG 304
Db 243 HKKFDALLTRMDFDEHKATSYERKGPDPFLDCVMENRDNSEGERLSTTNKALLNLEFTAG 302

Qy 305 TDSSTGTLEWLSLNHLNHPVLKARDELDTQVGDRLNLSDLPLKPLRYLKILETLRL 364
Db 303 TDTSSAIEWALAEWKNPALKKAQGEQVIGNNRRLLESDIPNULRAICKETFRK 362

Qy 365 YPPAPILPHVSSDITIEGNIIPRDTIIVINGMQRDPOLMNDATCFKPERF-----418
Db 363 HPSTPLNLPRISENPICVDGYITPNTLSVNIWAIGRDPVEMNPLEFYPERFLSRNS 422

Qy 419 --DVEGEEKLVAFGMGRACGFBPMQSVFTLGLLIQCDFWKRVSSE--KLDMTENN 475
Db 423 KIDPRGNDFFELIPFGAGRRICAGTRMGIVMVEYILGILVHSFDWKLPSVIELNMBEAF 482

Qy 476 ILSRLIPLEAMCKAR-----PLA 494
Db 483 LALQKAVEPLEAMVTPRLPIDVYAPLA 508

RESULT 8
US-08-606-505B-67

; Sequence 67, Application US/08606505B
; Patent No. 6114601
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yasuhiro
; APPLICANT: KIWAKAWA, Shigeto
; APPLICANT: SHIMADA, Yukihisa
; APPLICANT: OHBAYASHI, Masaya
; APPLICANT: SHIMADA, Ritsuko
; APPLICANT: OKINAKA, Yasushi
; TITLE OF INVENTION: NOVEL PLANT GENES
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112-3801
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
; COMPUTER: IBM PS/V
; OPERATING SYSTEM: MS-DOS Ver3.30
; SOFTWARE: PATENT AID Ver1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/606,505B
; FILING DATE: 23-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP44963/92
; FILING DATE: 02-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Perry, Lawrence S.
; REGISTRATION NUMBER: 31865
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-218-2100
; TELEFAX: 212-218-2200
; INFORMATION FOR SEQ ID NO: 67 :
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Campanula medium
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 180 to 1748
; IDENTIFICATION METHOD: by experiment
US-08-606-505B-67

Query Match 28.2%; Score 730; DB 3; Length 523;
Best Local Similarity 31.3%; Pred. No. 3.1e-59;
Matches 159; Conservative 117; Mismatches 208; Indels 24; Gaps 7;

Qy 4 VVSAYVFLVFLGVKVFQSRKLRNIPGPPPLPIIGNLNLEQIHRFFORMSKQYGN 63
Db 12 LVAAISLIATYSFIRFLPFSHHHLPPOPTGWPILGALPLGTGPHVSLADMAVKYGP 71

Qy 64 VVSLWFGSRLAVISSPTAYQBCFTKHDVALANRLPSLSKGKIFYNNTTVGSCSGHEHWR 123
Db 72 IMYKLGSGKTVASVSNPKAARAFKTHDANFSNRPIDGGPTYLAYNAQDMVFAEYGPWK 131

Qy 124 NLRRTALDVLSTQVHVSFGIRSDTKLMQRLVLAKNSNEEFARVEISSMFNDLTNN 183
Db 132 LLRLKLSLHMLGPKALELDWAVKVSFVGHMLKEMY--EQSSKSVPPVPPVPELTYAMAN 189

Qy 184 NINRMISGRKRYFGESEEMKNVEEA----REFRETVTETMLMGLMGLANKGDHLPFLRWFDQ 239
Db 190 MIGRIILTSRRFPVITSKLDSSASASVSEFQVMVLMRLMRMAGLFNIGDFIPIAMMDLQ 249

Qy 240 NVEKRLKISIKRYDSIILNKILHN--RASNDNRQNSMIDHLLKLQETQPOYYTDQI-----292
Db 250 GIORDMKVQKQKFDVLLNKNWIKETTESAHDKN---PDFLDILMAATQENTEGIQNLNLYN 307


```
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 510 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-948-564-4

Query Match          28.1%; Score 729; DB 3; Length 510;
Best Local Similarity 31.1%; Pred. No. 3.7e-59;
Matches 158; Conservative 116; Mismatches 192; Indels 42; Gaps 10;

QY  7 YAVLFLVFLGVKVF-----QSKRLRNIPPGPPPLPIIGNLNL--EQPIHFFQ 55
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 FSIYFITSILFIFFVFKLQVRSKTSCKLPGPRTPLIGNIHQIVGSLPVHYLK 70
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 56 RSKQYGNVSLWFGSLAVVISPTAYQECFTKHDVALANRLPSLSGKIFYFNNTVGS 115
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 NLADKYGLMHLKLGESNIIVTSPMAQIMTKTHDLNFDSDRDPFVLSRVSYNGSIVF 130
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 116 CSCEHWRNLRITALDVLSTQVHSPSGIRSDTKLMQRLVLAKNSNEEFARVEISS 175
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 131 SQHGVDYWRQKICTVELLTAKRVQSPRSIRREVEAEVLVK--IAATASEGGSIFNLQ 188
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 176 MFNDLTNNIMRLISGRFYGESEEMKNVEAREFRETVTMELEMLGLANKGDHLPFLRW 235
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 SIYSMTFG-----IAARAAGKKSRYQV-----FISNMHQLMLLGGFVADLYESSRV 238
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 236 FDFQNVKRLKISKRYDSILNKLHE---NRASNDQ---NSMIDHLLKLQETOPQYTT 289
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 FQMGATGKLEKVRHVTRDVLQDIIDEHKNRNSSEEREAEVLDVLLKFKQSEFRLT 298
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 290 DOIKGLALAMLPGGTDSTGTLEWSLSNLNHNPEVLKARDELDTQVQDRLNEDSLP 349
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 DDNIKAVIQDIFGGGTSSVWEGWMSLIRPNVMEEAQAEVRRVYDSKGVYDETELH 358
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 350 KLPYLRKIILETLRLYPAPILIPHVSSDITTEGFIIPRDTVIINGWGMQDQPOLWND 409
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 359 QLIYLSKIIKETWRLHPVLLVPRVSRRCQINGEIPSKTRIIINAWAIGRNPKYWG 418
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 410 ATCFKPERF-----DVEGEKKLVAFGMGRACPGEPMAQSVSFTLGLLIQCFDMK--- 461
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 419 TESFKPERFLNSSIDFRGTDPEFIPFGAGRRICGIFTFAIPNTIELPLAQLLYHFDWKL 478
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 462 RVSEKLDNTENNITLSR-----LIPJ 484
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 479 KMKNEELDMTESGITLRRQNDLCLIPJ 506
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 11
US-08-532-065B-2
; Sequence 2, Application US/08532065B
; Patent No. 5753507
; GENERAL INFORMATION:
; APPLICANT: Ohta, Daisaku
; APPLICANT: Mizutani, Masaharu
; TITLE OF INVENTION: Plant Geraniol/Nerol 10-Hydroxylase and
; TITLE OF INVENTION: DNA Coding Therefor
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 5753507artis Corporation
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: NJ
; COUNTRY: USA
; ZIP: 07936
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/532,065B
```

```
; FILING DATE: 22-SEP-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 495 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-532-065B-2

Query Match          28.0%; Score 725.5; DB 1; Length 495;
Best Local Similarity 34.6%; Pred. No. 7.4e-59;
Matches 165; Conservative 103; Mismatches 182; Indels 27; Gaps 9;

QY  1 LLAVVSYAVFLVFLVGLGVKFVQSRKLRNIPGPPPLPIIGNLNLLEQPIHFFQMSKQ 60
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  9 LLLLFCFLSCFLIPTTRSGRISRGATALPPGPPRLPIIGNIHLVGKHPHRSFAELSKT 68
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 61 YGNVSLWFGSLAVVISPTAYQECFTKHDVALANRLPSLSGKIFYFNNTT--VGSCSHG 119
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 YGPMVSLKLGSLNTVVIASPEAREVLRTHDQILSARPTNAVRSINHODASIVLWLPSSS 128
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 120 EHRNLRITALDVLSTQVHSPSGIRSDTKLMORLVLAKNSNEEFARVEISSMFDND 179
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 129 ARWRLRLSLVTQLLSQRIEATKALRMKNVKELVS--FISESSDREE--SVDISRVAFI 184
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 180 LYNNIMRMISGRF--YGESEEMKNVEAREFRETVTMELEMLGLANKGDHLPFLRWF 237
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 185 TTLNIIISLNFSDVGLSGNAKASINGVQ-----DTVISVMDAAGTPDAANTFFPFLRFLD 238
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 238 FQNVKRLKISKRYDSILN-----KILHENRASNDR---QNSMIDHLLKLQETOPQYTT 289
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 LQGVKTFKVCYTERLVRFVGRFIDAKIAEKSSQNNPKDYSKNDVFDNLLDYKGESELSI 298
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 290 DOIKGLALAMLPGGTDSTGTLEWSLSNLNHNPEVLKARDELDTQVQDRLNEDSLP 349
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 SD-IEHLLDMFTAGTDTSSSTLEWPMTELLKNPKTKMAKAQAEIDCVIGQNGIVEESDIS 357
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 350 KLPYLRKIILETLRLYPAPILIPHVSSDITTEGFIIPRDTVIINGWGMQDQPOLWND 409
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 358 KLPYLAQVVKETFRLLTPVPLLIIPKAESDABILGFNVLKDTQVLVNWVAIGRDPVSDWN 417
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 410 ATCFKPERF-----DVEGEKKLVAFGMGRACPGEPMAQSVSFTLGLLIQCFDMK 461
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 418 PSQFPERFLGKMDVGRDYEITLPPGAGRRICPGMPLAMKTVSLMLASLYSFDWK 474
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-09-126-420A-16
; Sequence 16, Application US/09126420A
; Patent No. 6376753
; GENERAL INFORMATION:
; APPLICANT: BATAED, YANNICK
; APPLICANT: ROBINEAU, TIBURCE
; APPLICANT: DURST, FRANCIS
; APPLICANT: WERCK-REICHT, DANIELE
; APPLICANT: DIDIERJEAN, LUC
; TITLE OF INVENTION: PURIFIED CYTOCHROME P450 CYP76B1 FROM HELIANTHUS
; TITLE OF INVENTION: TUBEROSUS AND ITS APPLICATIONS AS BIOCATALYST IN
; TITLE OF INVENTION: PARTICULAR FOR THE DEGRADATION OF ENVIRONMENTAL
; TITLE OF INVENTION: POLLUTANTS AND FOR ALTERING THE RESISTANCE OF PLANTS S
; FILE REFERENCE: 03715.0032
; CURRENT APPLICATION NUMBER: US/09/126,420A
; CURRENT FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/054,351
; PRIOR FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 27
```